Lee-Carter Mortality Forecasting, a Parallel GLM Approach, England & Wales Mortality Projections

A.E. Renshaw, S. Haberman

Actuarial Research Paper No. 140

January 2002

ISBN 1 901 615 63 4

"Any opinions expressed in this paper are my/our own and not necessarily those of my/our employer or anyone else I/we have discussed them with. You must not copy this paper or quote it without my/our permission".

Lee-Carter Mortality Forecasting, a Parallel GLM Approach, England & Wales Mortality Projections

A.E. Renshaw, S. Haberman

Abstract

The paper contains ways of re-interpreting the model structure underpinning the Lee-Carter mortality forecasting methodology, and proposes a parallel GLM based methodology. The gender and age specific England & Wales mortality experiences of the past half-century are analysed and the forecasts generated using both approaches compared.

Keywords: Mortality forecasting; Lee-Carter methodology; Generalised linear models; Mortality reduction factors

1. Introduction

A generalised linear modelling (GLM) regression based methodology for the construction of a class of temporal mortality reduction (adjustment) factors is described in Renshaw and Haberman (2000). Although initially designed with a view to the retrospective study of such factors, the methodology is also capable of a prospective interpretation and application, generating forecasts. Noting that this formulation also encompasses the model structure underpinning the Lee-Carter (LC) approach to mortality forecasting (Lee and Carter (1992), Lee (2000)), Renshaw and Haberman (2001) have investigated how the LC approach might be adapted to construct reduction factors for use in an actuarial context. In this paper, we investigate further how this GLM based approach to the construction of mortality reduction factors might be adapted along parallel lines to the LC approach for the purpose of mortality forecasting. A key difference between the two approaches centres on the interpretation of time, which, under the LC methodology is modelled as a factor and estimated by the singular value decomposition (SVD), and, under the GLM approach, is modelled as a covariate.

By way of illustration, mortality forecasts for both males and females, based on the England & Wales mortality experiences of the past half-century are made using both methods, and the results compared: the forecasts being of interest in their own right.

In Section 2, an outline of LC methodology is presented. This includes a discussion of the singular value decomposition (SVD) residuals, the graphical analysis of which would appear to have received little attention in this context. Lee (2000) contains a particularly informative resumé of the current state of LC methodology. In Section 3, a parallel GLM based methodology is described. Given that the basis of the formulation depends on modelling over-dispersed Poisson responses, it is possible to ensure that the actual total deaths equal the total expected deaths, for each relevant calendar year, as in the LC fitting process, by suitably augmenting the linear predictor. In Section 4, forecast mortality rates based on the England & Wales experience of the past half-century, using both methodologies are compared. This is conducted graphically. Model fitting is preceded by an exploratory analysis of trends in the crude mortality rates as an aid to choosing model structures, for the case of the GLM methodology. In Section 5, the forecasting of life expectancies at birth is discussed, again under both methodologies. In Section 6, the feasibility of extrapolating mortality rates beyond the available upper age limit is examined. Section 7 contains some concluding comments. There are two appendices, one containing the graphical exploratory analysis of log crude mortality

rates, the other including a discussion of issues concerning prediction errors under the GLM methodology.

2. The LC Mortality Forecasting Methodology

The Model:

The model underpinning the LC methodology (Lee and Carter (1992)), is the following:

$$\log m_{xt} = \alpha_x + \beta_x \kappa_t + \varepsilon_{xt} \qquad (2.1)$$

where

- m_{xt} denotes the central mortality rate for age x at time t
- α_x describes the shape of the age profile averaged over time
- β_x describes the pattern of deviations from the age profile
- κ_t describes the variation in the level of mortality with t,
- ε_{xt} denotes error.

The model is over-parameterised in the sense that the structure is invariant under either of the parameter transformations

$$\{\alpha_x, \beta_x, \kappa_t\} \mapsto \{\alpha_x, \beta_x / c, c\kappa_t\}$$

$$\{\alpha_x, \beta_x, \kappa_t\} \mapsto \{\alpha_x - c\beta_x, \beta_x, \kappa_t + c\}$$

for any constant c. Thus κ_t is determined up to a linear transformation, β_x is determined only up to a multiplicative constant, and α_x is determined only up to a linear adjustment.

Model Fitting:

It is envisaged that a rectangular array of crude mortality rates is available for fitting the model structure. Let (d_{xt}, e_{xt}) represent the data array, where

- d_{xt} denote the number of deaths at age x and time t, and
- e_{xt} denote the appropriate matching central exposures,

so that crude central mortality rates are given by $\hat{m}_{xt} = d_{xt} / e_{xt}$. For such an array, let

$$t = t_1,\ t_1 + 1,\ \ldots,\ t_1 + h - 1 = t_n;\ x = x_1, x_2, \ldots, x_k \text{ with } x_i < x_{i+1}$$

so that $h = t_n - t_1 + 1$ specifies the time range, and k specifies the number of age categories.

Under the LC methodology, the parameter estimates are normalised by stipulating that

$$\sum_{t=t_1}^{t_n} \kappa_t = 0 \text{ and } \sum_{all\,x} \beta_x = 1, \text{ so that } \alpha_x = \log \prod_{t=t_1}^{t_n} m_{xt}^{1/h} \qquad (2.2)$$

is a least squares error estimator. Allowing for constraints, the number of free parameters is $2k + h - 2$. The normalisation of β_x means that the resulting values indicate the relative rate of change of log mortality rates at different ages. Then the LC model fitting proceeds as follows:

- Estimate α_x as $\hat{\alpha}_x = \log \prod_{t=t_1}^{t_n} \hat{m}_{xt}^{1/h}$, the logarithm of the geometric mean of the crude mortality rates, averaged over all t, for each age x.

- Compute the matrix of statistics $[z_{xt}]=[\log \hat{m}_{xt}-\hat{\alpha}_x]$, and then estimate κ_t and β_x as the respective 1st right and 1st left singular vectors in the SVD of the matrix $[z_{xt}]$, subject to the above constraints.
- Finally, the estimated κ_t are adjusted such that the actual total observed deaths $\sum_{all\ x} d_{xt}$ equals the total expected deaths $\sum_{all\ x} e_{xt}\exp(\hat{\alpha}_x+\hat{\beta}_x\hat{\kappa}_t)$, $\forall t$.

It is possible to implement SVD using a variety of standard statistical packages offering this facility, such as GENSTAT (Payne *et al.* (1993)).

Residuals:

It would appear that little detailed attention has been given to the definition and analysis of residuals in the cited references. One possibility, implied in Tuljapurkar *et al.* (2000), is as follows. Denoting the (well ordered) singular values and the left and right singular vectors, generated under the SVD of the matrix $[z_{xt}]$ by s_i, $u_i(x)$ and $v_i(t)$ respectively, it follows that we can write

$$\log m_{xt}=\alpha_x+s_1u_1(x)v_1(t)+\sum_{i=2}^{r}s_iu_i(x)v_i(t) \qquad (2.3)$$

where $r=\min(h, k)$. Given that β_x and κ_t are estimated on the basis that

$$\beta_x\kappa_t=s_1u_1(x)v_1(t),$$

(subject to normalising constraints), comparison of (2.3) with (2.1) implies that the residual errors are represented as

$$\varepsilon_{xt}=\sum_{i=2}^{r}s_iu_i(x)v_i(t)=\log m_{xt}-\alpha_x-\beta_x\kappa_t. \qquad (2.4)$$

We refer to these as SVD residuals. Whereas it is possible to compute ε_{xt} either 'directly', using the first version of (4.2), or 'indirectly', using

$$\hat{\varepsilon}_{xt}=\log \hat{m}_{xt}-\hat{\alpha}_x-\hat{\beta}_x\hat{\kappa}_t,$$

only this latter 'indirect' method of computation is available when the adjustments, described above as part of the L-C estimating process, are made to κ_t, following the SVD of $[z_{xt}]$.

Alternatively, Renshaw and Haberman (2001) have used Pearson residuals, normally associated with a Poisson response model, namely

$$r_{xt}=\frac{d_{xt}-\hat{d}_{xt}}{\sqrt{\hat{d}_{xt}}},\ \hat{d}_{xt}=e_{xt}\exp\{\hat{\alpha}_x+\hat{\beta}_x\hat{\kappa}_t\} \qquad (2.5)$$

in the context of this model. It is of interest to note, that, while corresponding residual types differ in magnitude, the trivially established identity

$$\hat{\varepsilon}_{xt}=\log\frac{d_{xt}}{\hat{d}_{xt}},$$

crucially implies that they have identical signs.

Finally, either set of residuals may be standardised by dividing by $\sqrt{\phi}$, where the scale parameter ϕ is estimated according to

$$\hat{\phi}=\frac{\sum_{all\ x,t}\hat{\varepsilon}_{xt}^2}{v}\ \text{or}\ \hat{\phi}=\frac{\sum_{all\ x,t}r_{xt}^2}{v},$$

as the case may be, with $\nu = kh - 2k - h + 2 = (k-1)(h-2)$ degrees of freedom. This assumes there are no empty cells in the rectangular data array (with an obvious reduction, when this is not the case).

Model Forecasts:

Forecasts follow by modelling the values of κ_t as a time series using standard methods, with typically, the ARIMA(0,1,0) process, or random walk with drift, playing a central role in the approach. Denoting the resulting forecasts beyond the data time horizon as $\{\dot{\kappa}_{t_n+s} : s = 1,2,3,...\}$, forecast mortality rates are then computed as

$$\begin{aligned} \dot{m}_{x,t_n+s} &= \exp\{\log \hat{m}_{xt_n} + \hat{\beta}_x(\dot{\kappa}_{t_n+s} - \hat{\kappa}_{t_n})\} \\ &= \hat{m}_{xt_n} \exp \hat{\beta}_x(\dot{\kappa}_{t_n+s} - \hat{\kappa}_{t_n});\ s = 1,2,3, \ldots \end{aligned} \tag{2.6}$$

thereby ensuring the forecasts are aligned to the latest available crude mortality rates. If the latest data are judged to generate atypically shaped crude mortality rates (by age), it is possible to average across a few years at the end of the observed period.

For a discussion of the complexity and nature of some of the errors in such forecasts, see appendix B of Lee and Carter (1992).

Comment:

Whereas the method of model fitting is as described in the original paper by Lee and Carter (1992), the method of forecast alignment differs from that described in the original paper to reflect Lee's current thinking (Lee 2000).

3. A GLM Based Regression Approach to Mortality Forecasting

Mortality Reduction Factors:

Our approach to modelling is motivated by the concept of a mortality reduction factor, developed by UK actuaries (e.g. CMI Committee (1990)), and designed to be a means of adjusting current standard mortality tables, in order to predict future mortality trends.

In terms of μ_{xt}, the force of mortality at age x and time t, a mortality reduction factor (RF) is characterised by the equation

$$\mu_{xt} = \mu_{xt_0} RF(x,t),\ t \geq t_0 \tag{3.1}$$

subject to the constraint

$$RF(x,t_0) = 1\ \forall x. \tag{3.2}$$

The further constraints

$$1 \geq RF(x,t) > 0,\ \forall x,\ t > t_0,$$

plus decreasing monotonicity with increasing t, $\forall x$ and $t > t_0$, are implicit in the case of a *reduction* factor. It is envisaged that μ_{xt_0}, valid at time t_0, is completely specified, possibly in the form of a standard mortality table, or estimated, as described below. It is further envisaged, but not essential, that t_0 and μ_{xt_0} represent the current state in a dynamic process. The objective is to target RF by focusing on the same rectangular data array as above (Section 2), so that

$$t = t_1,\ t_1 + 1,\ \ldots,\ t_1 + h - 1 = t_n;\ x = x_1, x_2, \ldots, x_k \text{ with } x_i < x_{i+1},$$

but, this time, with a well defined time origin t_0, $t_1 < t_0 \leq t_n$.

To estimate μ_{xt_0}, consider the sub-interval $[t_{m_1}, t_{m_2}]$, such that $t_1 \le t_{m_1} \le t_0 \le t_{m_2} \le t_n$, with range $g = t_{m_2} - t_{m_1} + 1$. It is envisaged that the sub-interval is sited either at the extremity of the time interval so that $m_2 = n$ or is sited to coincide with a standard mortality table, such as an English Life Table, in which case $g = 3$. In general, g is small relative to h. Then μ_{xt_0} may be estimated as

$$\hat{\mu}_{xt_0} = \prod_{t=t_{m_1}}^{t_{m_2}} \hat{m}_{xt}^{1/g} \text{ or } \hat{\mu}_{xt_0} = \sum_{t=t_{m_1}}^{t_{m_2}} d_{xt} \Big/ \sum_{t=t_{m_1}}^{t_{m_2}} e_{xt} \qquad (3.3)$$

Note that the structure of the LC model, equation (2.1), (but subject to an additional constraint), is reproduced on taking logs of (3.1) and (3.2), and defining

$\alpha_x = \log \mu_{xt_0}$, $\log RF(x,t) = \beta_x \kappa_t$, s.t. $\kappa_{t_0} = 0$.

The Model:

In keeping with current UK actuarial practice when targeting the force of mortality (e.g. Forfar *et al.* (1988), Renshaw (1991), Renshaw *et al.* (1996)), we model the numbers of deaths D_{xt}, with matching central exposures e_{xt}, as the independent response variables of an over-dispersed Poisson GLM, with scale parameter ϕ, for which

$$E(D_{xt}) = e_{xt}\mu_{xt} = e_{xt}\mu_{xt_0} RF(x,t), \; Var(D_{xt}) = \phi E(D_{xt}).$$

Under the log link, the linear predictor

$$\eta_{xt} = \log e_{xt} + \log \mu_{xt_0} + \log RF(x,t), \text{ s.t. } \log RF(x,t_0) = 0$$

contains the offset

$$\log e_{xt} + \log \mu_{xt_0}.$$

Parameterised structures of potential interest, include

$$\log RF(x,t) = \beta_x (t - t_0) \qquad (3.4)$$

$$\log RF(x,t) = \beta_x (t - t_0) + \gamma_t (t - t_0) \qquad (3.4a)$$

$$\log RF(x,t) = \beta'_x (t - t_j)_- + \beta_x (t - t_0) \qquad (3.5)$$

$$\log RF(x,t) = \beta'_x (t - t_j)_- + \beta_x (t - t_0) + \gamma_t (t - t_0) \qquad (3.5a)$$

for $t \in [t_1, t_n]$, with

$$t_1 < t_j < t_0 \le t_n, \; (t)_- = t, \; t \le 0; \; (t)_- = 0, \; t > 0.$$

Thus (3.4) is representative of a pencil of lines (with a different slope parameter β_x for each age x), focused at t_0, and (3.5) is representative of a family of lines, hinged at t_j, also with a common focus at t_0. The extra term $\gamma_t (t - t_0)$, in the respective augmented versions (3.4a) and (3.5a), ensures that the adjustment made under the LC fitting process, whereby the total actual deaths equal the expected total deaths for each year is also accommodated under the GLM approach, while at the same time adhering to the constraint (3.2). In addition to satisfying this constraint, the specific form of this term is motivated by the close association between the independent Poisson response GLM and the analysis of contingency tables, which requires that the actual and expected column and row totals are identical under the independence hypothesis, the so-called Poisson trick (e.g. pp 248-249 Francis *et al.* (1993)). We note that the terms γ_t play no direct role in forecasting.

Model Fitting, Residuals:

Model fitting is conducted by optimising the over-dispersed Poisson quasi-likelihood. Under this formulation, and as an alternative to Pearson residuals (expression (2.5)), we have the choice of standardised deviance residuals

$$\frac{sign(d_{xt} - \hat{d}_{xt})\sqrt{dev(x,t)}}{\sqrt{\hat{\phi}}}$$

where, for an over-dispersed Poisson GLM

$$dev(x,t) = 2\{d_{xt} \log(d_{xt} / \hat{d}_{xt}) - (d_{xt} - \hat{d}_{xt})\}.$$

Similarly, the scale parameter is estimated as

$$\hat{\phi} = \frac{\sum_{all\ x,t} dev(x,t)}{\nu}$$

with ν degrees of freedom, determined, in part, by the degree of parameterisation in the model structure.

Model Forecasts:

Forecast mortality rates are again computed by alignment with the latest available crude rates, as under the LC methodology of Section 2, so that, for the parameterised structures represented by (3.4), (3.4a), (3.5) and (3.5a)

$$\dot{m}_{x,t_n+s} = \hat{m}_{xt_n} \exp \hat{\beta}_x s\,,\ s > 0. \qquad (3.6)$$

Comment:

By adopting the same method of forecast alignment, forecasts under the two methodologies may be assessed, relative to each other, by comparing the respective mortality reduction (adjustment) factors

$$\exp \hat{\beta}_x(\dot{\kappa}_{t_n+s} - \hat{\kappa}_{t_n})\,,\ s > 0;\ \text{and}\ \exp \hat{\beta}_x s\,,\ s > 0, \qquad (3.7)$$

based on the respective expressions (2.6) and (3.6).

4. Forecasting mortality rates for England & Wales

The Data:

The data comprise the number of deaths with matching central exposures to the risk of death as supplied by the Office for National Statistics. For each gender, these are cross classified by age, categorised in years as {<1, 1-4, 5-9, 10-14, … , 80-84, 85+}, and by individual calendar years, form 1950 to 1998 inclusive. Thus k =19 and h = 49.

Exploratory Data Analysis (EDA):

Time trends in the raw data, over the past half-century, are displayed by plotting the log crude mortality rates for each age separately in Appendix A. These plots are augmented by graphical displays of the first and second differences, of the log crude mortality rates. In addition, the first difference means, split at 1974, and the second difference (un-split) means, are superimposed on the respective graphs. Noteworthy features include the following:

- For males, there is a pronounced increase in the rate of mortality improvement, stemming from the 1970s, in the age range 40 to 80 years. This feature is less in evidence for females, being restricted to the age range 40 to 60.
- For each gender, there is a pronounced decrease in the rate at which mortality has been improving over the past quarter century, compared with the preceding quarter century, essentially in the age range 20 to 34 years. The increases in mortality rates observed during the 1990s at ages 25-34 for males and ages 20-24 for females are widely attributed to increases in the number of suicides and deaths related to HIV infection and AIDS (Daykin, (2000)).
- Outside the above age ranges, mortality trends are essentially linear (log scale) over the whole time span.

LC Modelling and Forecasting:

We follow the LC procedures exactly as described in Section 2, for each gender separately.

Although only the first right and left singular vectors, under the SVD of the matrix $[z_{xt}]$ generate the model structure, it is informative to view the first few such normalised vectors individually. We denote these as

$$\{\beta_i(x), \kappa_i(t) : i = 1,2,...,k\} \text{ subject to } \sum_{all\ x} \beta_i(x) = 1,\ \sum_{all\ t} \kappa_i(t) = 0,\ \forall i$$

where

$$\beta_i(x)\kappa_i(t) = s_i u_i(x) v_i(t)$$

with singular values $s_1 \geq s_2 \geq s_3 \geq ... \geq 0$ and

$$\beta_x = \beta_1(x), \kappa_t = \kappa_1(t)$$

(prior to the adjustment to κ_t), and the first five such sets of vectors are depicted in Fig 4.1(a & b). In addition to the well-defined structural patterns of the first SVD component, we also note that there are arguably well-defined patterns associated with the second component. However, on the basis of the proportion of the temporal variance in the transformed death rates explained by the first c components, measured by

$$100 \times \sum_{i=1}^{c} s_i^2 \Big/ \sum_{i=1}^{k} s_i^2 ,$$

the first component (c = 1) accounts for over 94% (Tuljapurkar *et al* (2000)). Details of the first five singular values, together with the cumulative percentages of the temporal variance explained, are as follows:

1st. five singular values

males: 7.51 (94.6%); 1.41 (97.9%); 0.56(98.4%); 0.42 (98.7%); 0.39 (99.0%)
females: 8.17 (94.7%); 1.42 (97.6%); 0.73(98.4%); 0.59 (98.8%); 0.47 (99.2%)

The modelling results are presented graphically in the various frames of Fig 4.2(a & b). Thus the SVD values of β_x and κ_t are depicted in the two uppermost LH frames, the latter after adjustment to ensure that the actual and expected total deaths are identical in each year. The magnitude of this adjustment may be ascertained by reference to the upper RH frame, in which the differences between the actual and expected total deaths, prior to adjustment, are depicted. In addition, the second RH frame supplements the matching LH frame by depicting the first differences in κ_t and

their split means. Thus for males, the pattern in the estimated κ_t essentially comprises two linear segments, hinged in the mid 1970s. For females however, the pattern in the estimated κ_t is essentially linear throughout the time span concerned. These findings, which are further supported by an analysis of the second differences in κ_t (not shown), are both consistent with the patterns in the two uppermost RH frames and much of the conclusions of the EDA described above.

The standardised SVD residuals are depicted in the third pair of frames, comprising 442 positive and 489 negative residuals for males, and 454 positive and 477 negative residuals for females. We remark that a plausible explanation of the somewhat unsatisfactory nature of the residual plot against calendar year for males is provided shortly (Fig 4.4a). We also note that the apparent discontinuity in these plots for both males, and, to a greater extent for females, between the adjacent years 1963 and 1964, (a feature which also surfaces in Fig 5.1), would appear to be intrinsic to the data.

The forecasts (and prediction limits), also featured in the second LH frame, are generated using the ARIMA(1,1,0) process, which is found to provide the most satisfactory fit to the κ_t,

$$\nabla\kappa_t - \lambda = \varphi(\nabla\kappa_{t-1} - \lambda) + \varepsilon_t \qquad (4.1)$$

where ∇ is the differencing operator and $\{\varepsilon_t\}$ denotes white noise. Such processes are selected and fitted using standard time series methodology (e.g. Hamilton (1994)) and then used to generate forecast values. Details are as follows:

ARIMA(1,1,0) parameter estimates (standard errors)

	$\hat{\varphi}$	$\hat{\lambda}$
males:	-0.532 (0.128)	-0.3041 (0.0629)
females:	-0.572 (0.122)	-0.3525 (0.0725)

For comparison purposes, we note that Tuljapurkar *et al.* (2000) using the LC model with $\varphi = 0$, so that an ARIMA (0,1,0) process is fitted, obtain a parameter estimate of $\hat{\lambda} = -0.3427$ for UK males and females combined, based on data for 1950-94, which is comparable to the above estimates. Renshaw and Haberman (2001) in modelling data on UK male annuitants and pensioners at ages 60-100 (provided by the CMI Bureau) fit a range of LC models and obtain λ estimates as follows:

Annuitants 1946-94, ARIMA (0,1,1)

$\hat{\lambda} = -0.457$

Pensioners 1983-94, ARIMA (0,1,0)

$\hat{\lambda}$ in the range –0.770 to –0.889 depending on the model used.

These estimates are dramatically larger in magnitude than those for the national population and highlight the pace of mortality decline and the importance of the selection processes operating in these sub-populations.

In addition, the temporal profiles of the product $\hat{\beta}_x\tilde{\kappa}_t$ for each age x are depicted in the lower LH frames of Fig 4.2(a & b), where

$\tilde{\kappa}_t = \hat{\kappa}_t$, for $t_1 \le t \le t_n$; and $\tilde{\kappa}_t = \dot{\kappa}_t$, for $t > t_n$

and their translations $\hat{\beta}_x(\tilde{\kappa}_t - \hat{\kappa}_{t_n})$ in the lower RH frames. Thus for $t > t_n$, the lower RH profiles depict the log mortality reduction factors for each age x.

GLM Modelling and Forecasting:

The offset term $\log \mu_{xt_0}$ is computed first using either version of (3.3), with reference to the latest available English Life Table ELT15 (Registrar General (1997)), centred on $t_0 = 1991$, and based on the data of the triennium 1990-92: the choice of formula being of negligible consequence in this context. For males, we illustrate in detail the hinged line mortality reduction factor structure, focused at t_0, equation (3.5a), siting the hinge t_j at 1975, in accordance with the earlier findings of the exploratory data analysis and the following deviance profile generated by re-siting the hinge as follows

Deviance profile, hinge sited in the year t_j

year t_j	1970	1972	1974	1975	1976	1978	1980
deviance	9674	9472	9384	9380	9389	9433	9596

For females, in accordance with the bulk of the findings in the EDA, we illustrate the pencil of lines, focused at t_0, equation (3.4a). In both cases, the adjustment, which equates actual and expected total deaths annually, is applied.

The results are again presented in graphical form. Thus the estimated values of β_x and the differences between the actual and expected total deaths, prior to adjustment, are depicted in the upper frames of Fig 4.3(a & b). Then, since κ_t has been replaced by a covariate t (with or without a hinge), there is no equivalent to the second pair of frames in Fig 4.2 (a & b) under the current formulation. Consequently, the pair of standardised deviance residuals plots is depicted next, comprising 426 positive and 505 negative residuals for males, and 434 positive and 497 negative residuals for females. Finally the temporal profiles

$$\log RF(x,t) = \hat{\beta}_x(t-t_0) + \hat{\gamma}_t(t-t_0),\ t_1 \le t \le t_n, \text{ or}$$

$$\log RF(x,t) = \hat{\beta}'_x(t-t_j)_- + \hat{\beta}_x(t-t_0) + \hat{\gamma}_t(t-t_0),\ t_1 \le t \le t_n,$$

as the case may be, and

$$\log RF(x,t) = \log RF(x,t_n) + \hat{\beta}_x t,\ t > t_n$$

for the log reduction factor, estimated and projected, for each age x, are depicted in the lower LH frames, together with their translation

$$\log RF(x,t) - \log RF(x,t_n)$$

in the lower RH frames.

Comparison of Results:

As noted in Section 3, by adopting the same method of alignment, forecasts under the two methodologies may be compared by means of their respective mortality reduction (adjustment) factors (3.7), which we re-express as

$$\log RF_{LC}(x,t_n+s) = \hat{\beta}_x(\hat{\kappa}_{t_n+s} - \hat{\kappa}_{t_n}),\ s > 0 \qquad (4.2a)$$

$$\log RF_{GLM}(x,t_n+s) = \hat{\beta}_x s,\ s > 0 \qquad (4.2b)$$

and which are depicted as part of the respective lower RH frames in Figs 4.2 & 4.3.

We remark that for an ARIMA(0,1,0) process, (equation (4.1) with φ preset to zero), which has been used for forecasting US mortality (Lee and Carter (1992) and mortality in G7 countries (Tuljapurkar *et al.* (2000)), equation (4.2a) reduces to

$$\log RF_{LC}(x,t_n+s) = \hat{\beta}_x \lambda s,\ s > 0,$$

the same form as (4.2b), while under an ARIMA (1,1,0) process, as used here, this expression represents the asymptotic form of (4.2a). Thus under the LC approach as here, the asymptotic form of the predictions in the lower RH frame of Fig 4.2 (a & b) is derived from the β_x displayed in the corresponding upper LH frame by multiplying by the (negative valued) general mean λ of the particular ARIMA(1,1,0) process concerned. On the other hand, under the GLM approached, the gradients of the predictions in the lower RH frame of Fig 4.3 (a & b) are as displayed in the corresponding upper LH frames. Hence the upper LH frame in Fig 4.2a might be loosely interpreted as a scaled mirror image of the upper LH frame in Fig 4.3a.

For males, a comparison of the projected portions of the lower RH frames in Fig (4.2a) and Fig (4.3a) indicates a different pattern of forecasts, across ages, under the two methodologies. Thus, the LC approach is based on a multiplicative decomposition of time effects which allows for their collective modelling across all ages and results in a more focused pencil of predictions than under the GLM approach. The GLM approach is designed to capture and project the more recent age specific time trends in mortality, depending crucially on the siting of the hinge (1975) and origin (1991). The likely source for these differences would appear to depend on two inter-related features involving the apparent lack of response in the time series forecasts to the significant hinge feature in the estimated κ_t (second RH frame in Fig 4.2a), coupled with a degree of inflexibility induced by the collective modelling of time effects across all ages, which may, or may not, be judged to be appropriate. Thus, for example, under an ARIMA(0,1,0) process, forecasts are generated by projecting the straight line joining the first and final estimated values of κ_t, irrespective of the intermediate values and hence of any hinge effect, which is then transmitted to each age x on multiplying κ_t by β_x, producing effects similar to those depicted in the lower frames of Fig 4.2a. On the other hand, under the GLM approach using the linear predictor (3.5a), forecasts are generated by projecting the more flexible age specific 'hinged lines' of best fit that have been perturbed to ensure that total actual and total expected annual deaths are identical (within the fitting period). Further evidence for this is to be found in the respective residual plots against time (the penultimate RH frames in Figs 4.2a & 4.3a), which are elaborated upon in Fig 4.4 (a & b) by plotting the residuals against time for each age separately. In particular, while there are quite distinctive U-shaped patterns in the age bands <1, 1-4, 25-29, 30-34 in Fig 4.4a under the LC methodology, these patterns are largely dispersed in Fig 4.4b under the GLM methodology.

Specifically for the male experience, reference to the upper LH and lower RH frames in Fig 4.3a indicates a small predicted increase in mortality rates in the age band 25-29 accompanied by no change in predicted mortality in the adjacent age band 30-34, reflecting the recent trends identified earlier as being attributable to deaths for suicides or HIV and AIDS related causes. A cross-reference to the upper LH and lower RH frames in Fig 4.2a, indicates that this would appear to be an important consideration overlooked under the LC approach. Another different feature is the more rapidly decreasing mortality rate projections at the very youngest ages, below one year of age in particular, under the GLM approach.

For females, comparison of the respective upper LH and lower RH frames of Fig (4.2b) and Fig (4.3b) indicates very similar patterns of forecasts by age. However, a somewhat less focused pencil of predictions does result when the hinged linear predictor (3.4a) is utilised, thereby capturing some of the finer detail identified

in the exploratory data analysis. Details are omitted but are available from the authors.

One clear difference between the two methodologies concerns the different time spans used in the estimation of α_x, or equivalently $\log \mu_{xt_0}$, as the case may be. Thus, while all the annual crude mortality rates covering the compete time span 1950-98 were used in the estimation of α_x under the LC methodology, only the annual crude mortality rates of the triennium 1990-92 were used in the estimation of $\log \mu_{xt_0}$ under the GLM methodology. To investigate this and other associated issues for the data sets in question, it is possible to reverse the roles of these different time spans and to then re-fit the models accordingly. Thus, in the upper two frames of Fig 4.5, the projections obtained for the respective male and female experiences under the LC methodology are displayed, using only the annual crude mortality rates of the triennium 1990-92 to estimate α_x in the first instance. The similarity of these diagrams with their respective counterparts, the lower RH frames of Fig 4.2 (a & b), is noteworthy. Then in the lower two frames of Fig 4.5 the projections using the perturbed hinged line (HL) and perturbed lines (L) for the respective male and female experiences are displayed, on the basis that the whole of the available time span was used in the estimation of the offset term $\log \mu_{xt_0}$ (using the first version of (3.3)) while siting the origin in 1974. We recall that this point is immediately adjacent to the location of the hinge for the male experience. Comparison with the respective counterparts in Fig 4.3 (a & b) reveals major differences for the male experience, thereby demonstrating the critical role played by the relative positioning of the hinge and origin, while relatively modest local differences across ages are revealed for the female experience.

5. Forecasting Life Expectancy at Birth

Two different approaches are of interest for forecasting life expectancies on a period basis: one involving the direct computation of life expectancies from forecast mortality rates, the other involving the time series analysis of the annual life expectancies at age x $\{e_t(x): t = t_1, t_1+1, ..., t_n\}$, as computed using crude mortality rates, to generate forecasts directly. While the first approach leads to compatible mortality rate and life expectancy forecasts, the second approach does not.

Both approaches are illustrated in Fig 5.1 for the case $x = 0$: life expectancies at birth. Thus, under the second approach, annual life expectancies for 1950 to 1998 are displayed, together with the forecast life expectancies at birth and their limits, generated by an ARIMA(1,1,0) process

$$\nabla e_t(0) - \lambda = \varphi(\nabla e_{t-1}(0) - \lambda) + \varepsilon_t,$$

operating on the annual life expectancies, for which

ARIMA(1,1,0) parameter estimates (standard errors)

	$\hat{\varphi}$	$\hat{\lambda}$
males:	-0.460 (0.133)	0.1841 (0.0277)
females:	-0.568 (0.123)	0.1818 (0.0407)

In addition, forecast life expectancies, computed directly from the respective forecast mortality rates arising from both the fitted LC and fitted GLM models of Section 4 have been superimposed for comparison.

For the female experience, in which log-linearity over time is a dominating influence (as discussed in Section 4), the forecasts based on both the LC and GLM models are, for all practical purposes, identical, with both forecasts understating the equivalent time-series based forecast. For the male experience, higher life expectancies are forecast under the GLM model than under the LC model, reflecting the differences in the model structures that generate the respective forecast mortality rates. Here, the time series forecast is intermediate between both these forecasts. (By switching from the perturbed hinged-line structure (3.5a) to the perturbed linear structure (3.4a), the resulting life expectancy forecast also understates the time-series based forecast, being slightly more conservative than the LC forecast). The conservative nature of the life expectancy forecasts under the LC approach, compared with the forecasts under the direct time-series approach (both genders), is consistent with the findings of Lee and Carter (1992) in their forecasting of US life expectancies.

6. An Extreme Age "Closing Out" Procedure

As noted by Lee and Carter (1992) and Lee (2000), a widely accepted procedure in demography for "closing out" model life tables (by age), in which observed mortality rates at extreme ages 85+ are replaced by a sequence of extrapolated mortality rates at ages 85-89, 90-94, ...,105-109 is the method of Coale and Guo (1989). The procedure uses the assumption of a steady decrease, rather than Gompertzian constancy, in the rate of increase in mortality with ages above 75. Thus, for the format of abridged age specific central mortality rates, which is of immediate interest in relation to the national data for England and Wales, the Coale-Guo assumption requires that

$$\log({}_5m_{80}) - \log({}_5m_{75}) = k$$

$$\log({}_5m_{85}) - \log({}_5m_{80}) = k - R$$

$$\vdots \qquad \vdots \qquad \vdots$$

$$\log({}_5m_{105}) - \log({}_5m_{100}) = k - 5R$$

which, on summing, imply

$$\log({}_5m_{105}) - \log({}_5m_{75}) = 6k - 15R.$$

Then, given a set of central mortality rates up to and including ${}_5m_{80}$, by additionally assigning

$${}_5m_{105} - {}_5m_{75} = 0.66 \qquad (6.1)$$

as suggested in Coale and Guo (1989), the above equations uniquely determine k, R and ${}_5m_{85}, {}_5m_{90}, \ldots, {}_5m_{105}$.

Empirical evidence supportive of the Coale-Guo assumption for the England and Wales national experience is provided by the graphical analysis of the crude mortality rates, Fig 6.1, for the triennium 1990-92 which were used in the construction of ELT15 (Registrar General 1997). Thus, in the upper and central LH frames of Fig 6.1, log crude mortality rates, by individual year of age (75 to 94 inclusive), are plotted against age, for the respective genders. The matching residual plots, displayed as the RH frames in Fig 6.1, are derived as a result of fitting quadratic curves (also displayed in the RH frames), using the method of ordinary least squares. Details of the fitted curves

$$\log \hat{m}_x = a + bx + cx^2 + error$$

are as follows

	Parameter estimates with (standard errors)	
	male experience	*female experience*
a	-11.77 (0.930)	-13.31 (0.424)
b	0.149 (0.022)	0.158 (0.010)
c	-0.000381 (0.000131)	-0.000338 (0.0000595)

in which the statistically significant, negative valued parameters *c*, are supportive of the underpinning Coale-Guo assumption.

This graphical analysis is supplemented by the two lower frames in Fig 6.1, in which the respective first differences of the log crude mortality rates for each gender are displayed, together with least squares fitted lines, both horizontal (for implied Gompertzian constancy) and general (consistent with the Coale-Guo assumption). Details of the fitted lines

$$\log \hat{m}_{x+1} - \log \hat{m}_x = A + Bx + error$$

are as follows

	Parameter estimates with (standard errors)	
	male experience	*female experience*
A	0.168 (0.0594)	0.154 (0.0306)
B	-0.000991 (0.000702)	-0.000633 (0.000361)

in which the negative valued slope parameters *B*, are supportive of the underpinning Coale-Guo assumption.

In addition, it is possible to monitor the suitability of the constraint (6.1), by noting that the separate curve and line fitting exercises imply that we can write either

$$\log m_{x+1} - \log m_x = (b + c) + 2cx \qquad (6.2)$$

or

$$\log m_{x+1} - \log m_x = A + Bx, \qquad (6.3)$$

as the case may be. Then, typically (6.3), say, implies that

$$\log m_{x+5} - \log m_x = 5A + 10B + 5Bx$$

which, subject to the approximation ${}_5m_x = \prod_{i=0}^{4} m_{x+i}^{1/5}$ implies that

$$\log {}_5m_{x+5} - \log {}_5m_x = 5A + 20B + 5Bx$$

and which, in turn, implies that

$${}_5m_{105} - {}_5m_{75} = {}_5m_{75}\{\exp(30A + 2745B) - 1\};$$

with an obvious equivalent expression based on (6.2). The values of ${}_5m_{105} - {}_5m_{75}$ generated by these methods are as follows

	Values of ${}_5m_{105} - {}_5m_{75}$	
	male experience	*female experience*
fitted quadratic	0.675	0.676
fitted line	0.654	0.665

We note that this analysis is sensitive to both the quality of the available data at elderly ages and the attendant potential for outliers. However, we conclude that this simple analysis, together with unreported similar analyses of equivalent data sets used in the construction of certain previous ELTs, is generally supportive of the conditions underpinning the Coale-Guo extrapolation technique.

By applying the procedure for each calendar year, thereby projecting (age-wise) both observed and forecast mortality rates in the age range 80-84, the technique somewhat arbitrarily emphasises the importance of the mortality pattern in the age range 80-84 by translating it to higher age groupings, while disregarding the available data for the age band 85+. On this basis we have recomputed the life expectancies at birth for both genders and both modelling approaches, to obtain the following results:

Life expectancy at birth: 2020 projections

	Male experience			*Female experience*		
	LC	*GLM*	*diff'ce*	*LC*	*GLM*	*diff'ce*
without C-G	78.21	79.83	-1.62	83.40	83.43	-0.03
with C-G	77.81	78.79	-1.28	82.12	82.12	0.00
difference	0.70	1.04		1.28	1.31	

To compute these values, we have found it necessary to interpolate linearly ${}_5m_{80}$ and the Coale-Gue projections ${}_5m_{85}, {}_5m_{90}, \ldots$ prior to mapping the force of mortality to the probability of death (equation (B.1) Appendix B). On this basis, the net effect of applying the Coale-Guo extrapolation procedure in these experiences is to reduce the forecast life expectancies at birth by the amounts shown.

As we note from Fig 4.2 (a & b) and Fig 4.3 (a & b) (top LH frames), the estimated $\hat{\beta}_x$ values are different at ages 75-79, 80-84 and 85 and over, implying a more rapid decrease in mortality over time at ages over 85 than ages 75 and 84. With the Coale-Guo technique, the mortality rate for the 85 and over age group is replaced by an estimated sequence of ${}_5m_y$ for y = 85, 90,...,105 whose values we effectively determined by ${}_5m_{75}$ and ${}_5m_{80}$ (as described above). This has the effect of reducing the impact of the downward trends in mortality at these advanced ages and hence leads to the reduced forecasts of life expectancy at birth, shown by the numerical estimated above.

In the recent literature, Lindbergson (2001) has suggested a Makeham model for mortality in Sweden, with the exponential growth with age replaced by a linear term at very high ages, *viz*

$$\mu_x = a + b(x - \omega_0) \text{ for } x > \omega_0 .$$

The relative effectiveness of this approach and the established method of Coale and Guo (1989) remains to be tested.

7. Concluding Comments

- The use of a GLM regression based methodology for forecasting mortality rates is perhaps unusual and is suggested on purely pragmatic grounds, given the similarities in structure to the LC time-series methodology. As stated in the introduction, the idea has its roots in the retrospective study of UK actuarial mortality reduction factors (Renshaw and Haberman (2000)).
- Both the GLM and LC approaches attempt to capture underlying trends in the raw data and to extrapolate these, thereby generating plausible mortality projections under the crucial assumption that existing trends continue into the future. In this respect, we identify closely with the sentiment (not always practised), expressing a preference for the term 'projection' over the term 'forecast' in Section 9.29, pp 211-12 of Benjamin and Pollard (1993), on the grounds that "there is no implication that [established trends] will be maintained".

- The status of α_x, describing the shape of the base age profile averaged over time, is similar under the two approaches. It is subtracted from the log crude mortality rates as a first stage in the fitting process under the LC methodology and subtracted from the linear predictor as an integral part of the fitting process under the GLM methodology. There is a wide choice of possible time bases on which to construct the estimates.
- The key difference between the methodologies concerns the treatment of time. Under the LC approach, time is modelled as a multiplicative factor and estimated, together with the β_x, by SVD. Under the GLM approach, time is modelled as a known covariate, having first established a well-defined origin, and only the β_x are estimated accordingly. (Under the LC approach, a time origin is generated naturally as a consequence of the normalising constraints (2.2), but unlike the GLM approach, no specific use is made of its exact location).
- It is possible to ensure that the total actual deaths and the total expected deaths match for each year in both approaches. This is done without compensating adjustments to the β_x under the LC approach, but with compensating adjustments to the β_x under the GLM integrated fitting approach.
- Forecasts are generated by extrapolated, age specific, straight lines (on the log scale), either exactly (under the GLM or LC approach in combination with an ARIMA(0,1,0) process, as used exclusively in Lee and Carter (1992), Lee (2000) and Tuljapurkar *et al.* (2000)), or asymptotically (under the LC approach in combination with an ARIMA(1,1,0) process, as discussed in Section 4).
- There is a possible concern, with the GLM approach, as expressed in Section 4.4.6 pp 122-24 of McCullagh and Nelder (1989), that such predictions might be sensitive to the choice of link function, but this has been allayed, in the current context of projecting mortality rates, by a practical investigation conducted in Renshaw and Haberman (2000).
- On the basis that, under the GLM approach, time is modelled as a known covariate and α_x is modelled as a known offset, the only source of error generated, from within the model, under the fitting process, involves the $\hat{\beta}_x s$. On this basis, as described in Renshaw and Haberman (2000), it is possible to construct the prediction limits for $\log RF$ quoted in Appendix B, which readily map into prediction limits for *RF*. Also under this limited error perspective, it is possible to construct approximate prediction limits for life expectancy as described in Appendix B. A realistic criticism of such limits however, concerns their relative narrowness. A possible way of increasing such limits might be by incorporating the error in $\hat{m}_{xt_n}$, used in the alignment of forecasts, as defined when stipulating the variance of the response in specifying the model. Just how this might be achieved remains a challenge. Issues concerning the construction of prediction limits for the LC method are discussed in Appendix B of Lee and Carter (1992).
- There is evidence from the male mortality study reported in Section 4 that the GLM approach is more successful at capturing age specific mortality trends in the data, a feature intrinsic to the respective approaches.

- Different approaches to forecasting life expectancy are described and illustrated for $e_t(0)$. Direct modelling of $e_t(0)$ using time series methods is compared with methods based on forecasting the underlying mortality rates by either LC or GLM approaches. While an improvement in life expectancy at birth is forecast, for each gender, the established gap between female and male life expectancy at birth is forecast to decrease under the GLM approach.

Acknowledgements

The authors wish to thank the CMI Bureau of the Institute & Faculty of Actuaries for providing financial support. We are also indebted to R.D. Lee for kindly responding to certain enquiries.

References

Benjamin, B. and J.H. Pollard (1993). *The Analysis of Mortality and Other Actuarial Statistics.* Institute and Faculty of Actuaries, Oxford.

CMI Committee (1990). Standard tables of mortality based on the 1979-82 experiences. *Continuous Mortality Investigation Report*, Institute and Faculty of Actuaries, 10, 50-53.

Coale, A. and G. Guo (1989). Revised regional model life tables at very low levels of mortality. *Population Index, 55, 613-643.*

Daykin, C.D. (2000). The recent trend of mortality in the United Kingdom. *British Actuarial Journal,* 6, 861-871.

Francis, B., M. Green and C. Payne (1993). *The Glim System: Release 4 Manual.* Clarendon Press, Oxford.

Forfar, D.O., J.J. McCutcheon and A.D. Wilkie (1988). On graduation by mathematical formula. *Journal of the Institute of Actuaries*, 115, 1-135.

Hamilton, J.D. (1994). *Time series analysis*. Princeton University Press, Princeton, New Jersey.

Lee, R.D. (2000). The Lee-Carter method of forecasting mortality, with various extensions and applications. *North American Actuarial Journal* (with discussion), 4, no 1, 80-93.

Lee, R.D. and L. Carter (1992). Modelling and forecasting the time series of U.S. mortality. *Journal of the American Statistical Association* (with discussion), 87, 659-71.

Lindbergson M. (2001). Mortality among the elderly in Sweden 1988-87. *Scandinavian Actuarial Journal,* 1, 79-94.

McCullagh, P. and J.A. Nelder (1989). *Generalised Linear Models.* Chapman & Hall.

Payne, R.W. *et al.* (1993) *Genstat 5 Release 3 Reference 4 Manual.* Clarendon Press Oxford.

Registrar General (1987). *English Life Tables No 14.* Office of National Statistics, Series DS No 14, HMSO, London.

Registrar General (1997). *English Life Tables No 15.* Office of National Statistics, Series DS No 14, HMSO, London.

Renshaw, A.E. (1991). Actuarial graduation practice and generalised linear and non-linear models. *Journal of the Institute of Actuaries*, 118, 295-312.

Renshaw, A.E. and S. Haberman (2000). Modelling for mortality reduction factors. *Actuarial Research Paper No. 127, City University, London.* (under review).

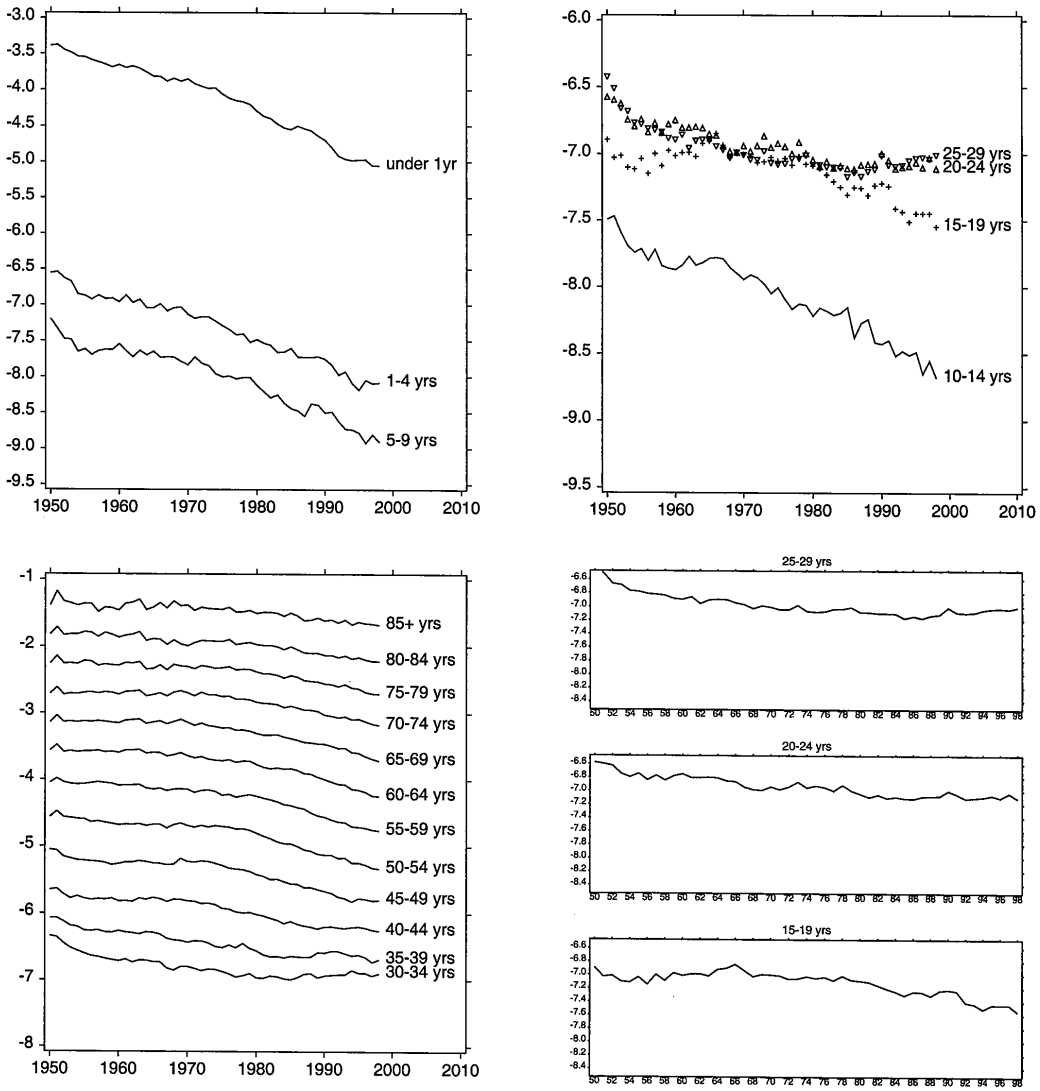
Renshaw, A.E. and S. Haberman (2001). On the forecasting of mortality reduction factors. *Actuarial Research Paper No. 135, City University, London.* (under review).

Renshaw, A.E., S. Haberman and P. Hatzopoulos (1996). Recent mortality trends in U.K. male assured lives. *British Actuarial Journal,* 2, 449-477.

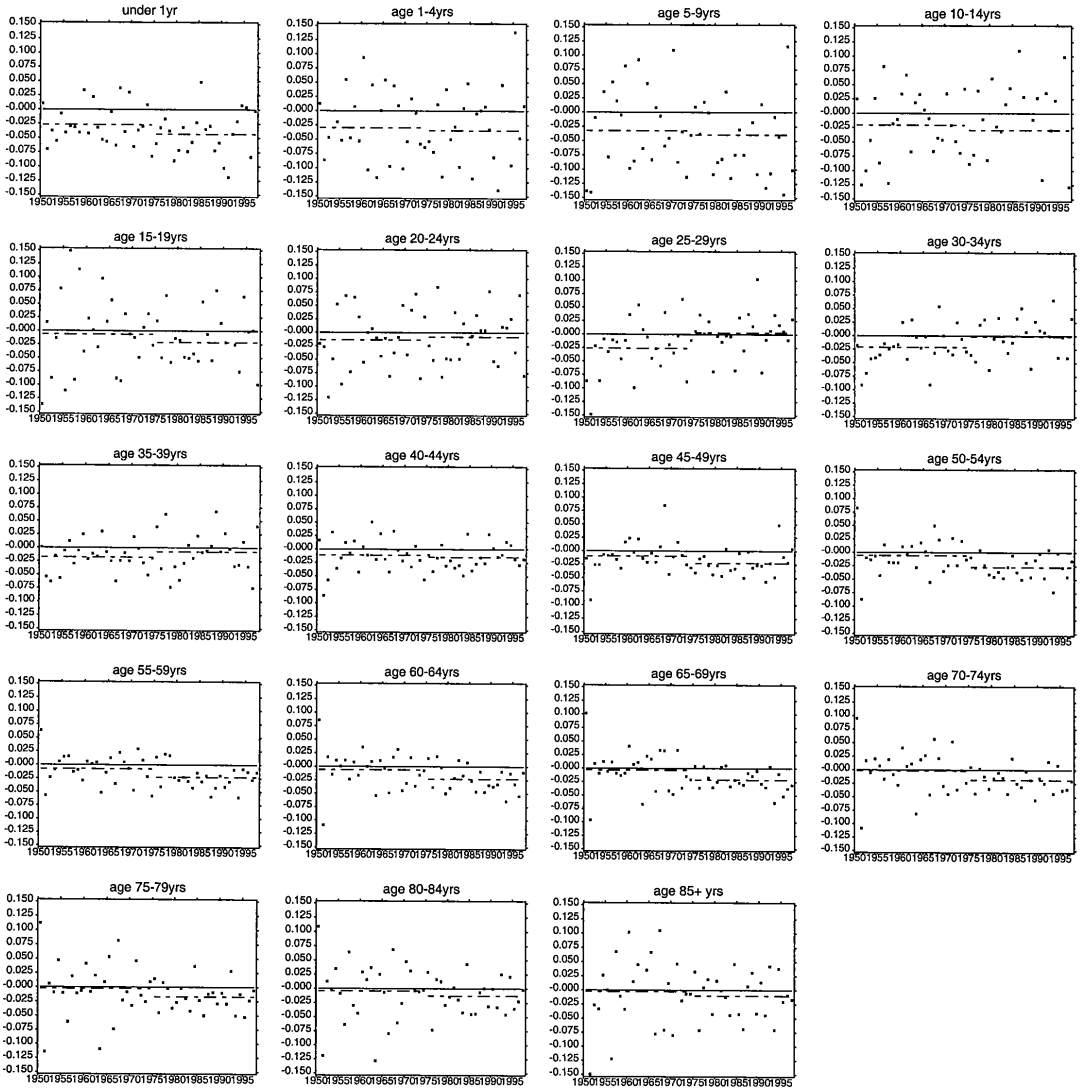
Tuljapurkar, S., N. Li and C. Boe (2000). A universal pattern of mortality decline in the G7 countries. *Nature,* 405, 789-792.

Appendix A
(Graphical analysis: log crude mortality rates by year, for age and each gender)

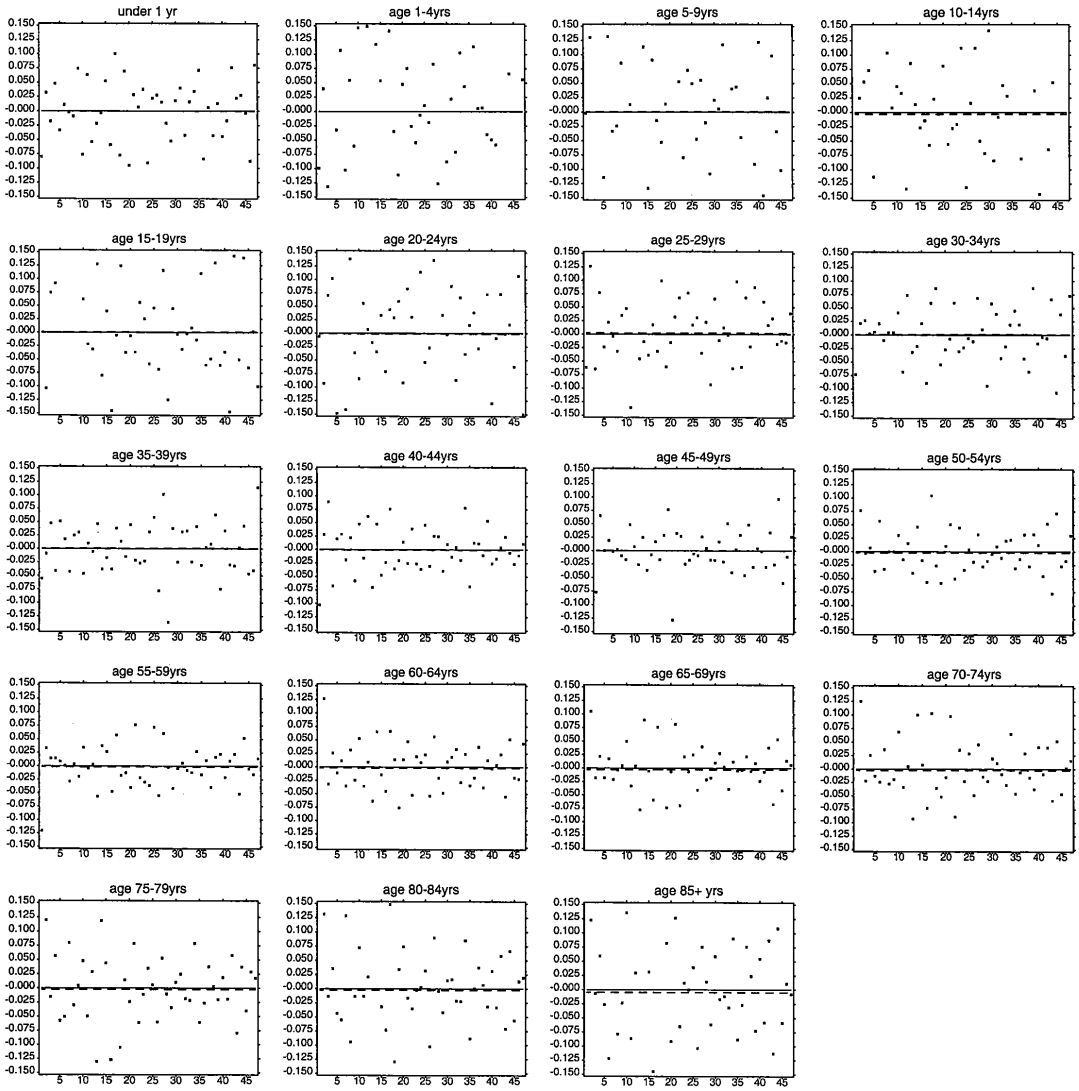
List of figures



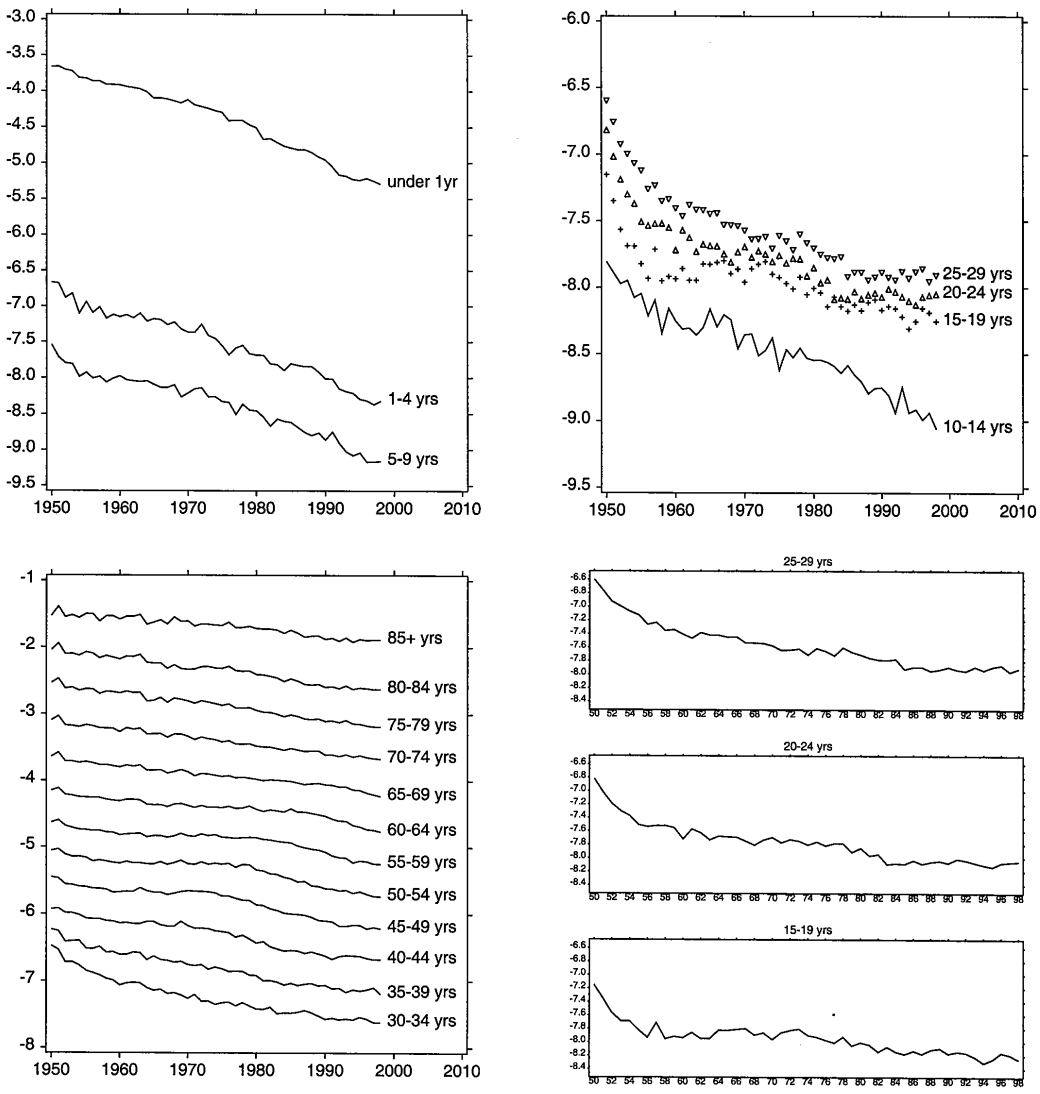
A.1 Male log crude mortality rates, by year, for age



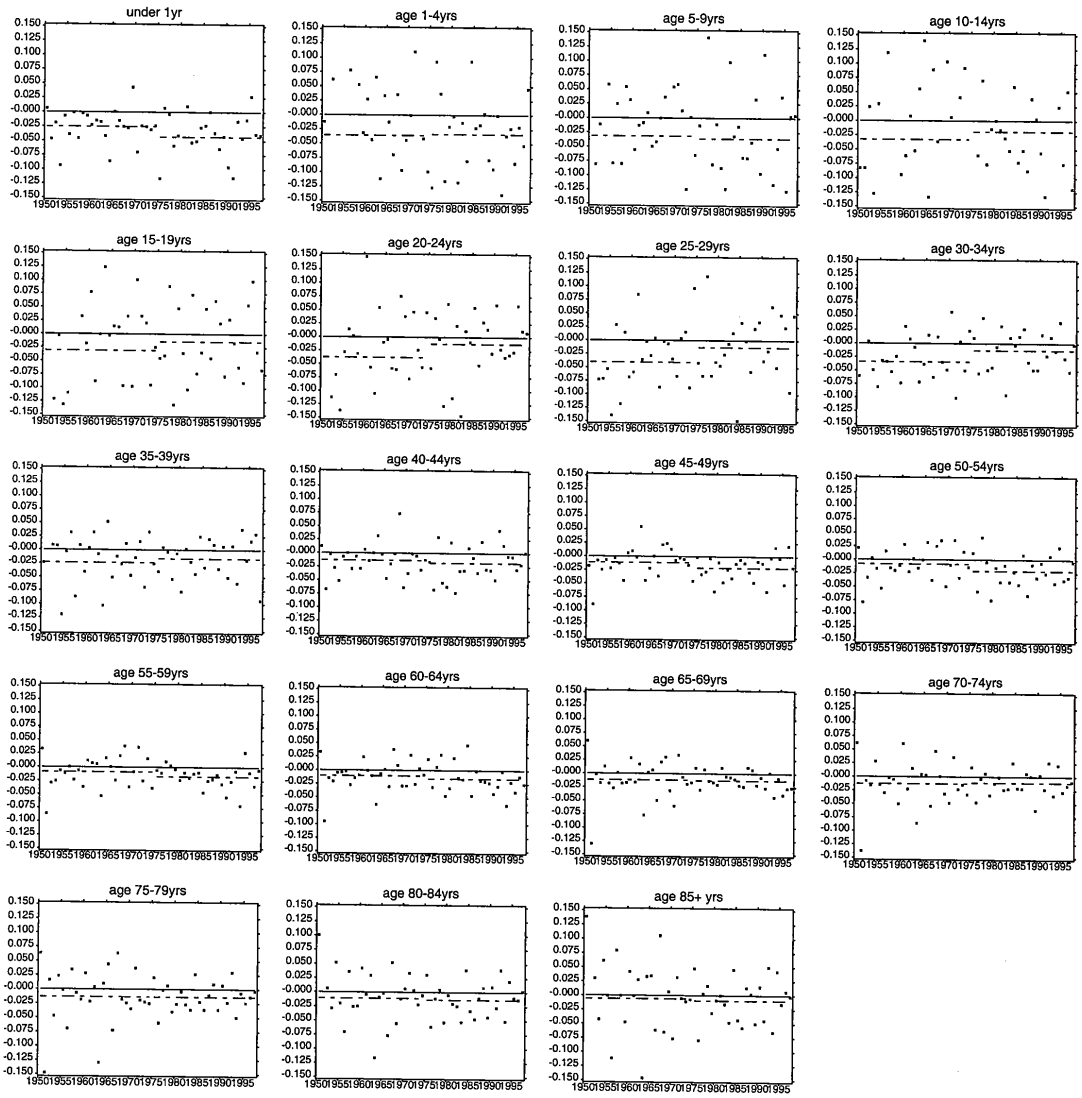
ɪrtality rate 1st differences (with split means), by year, for age



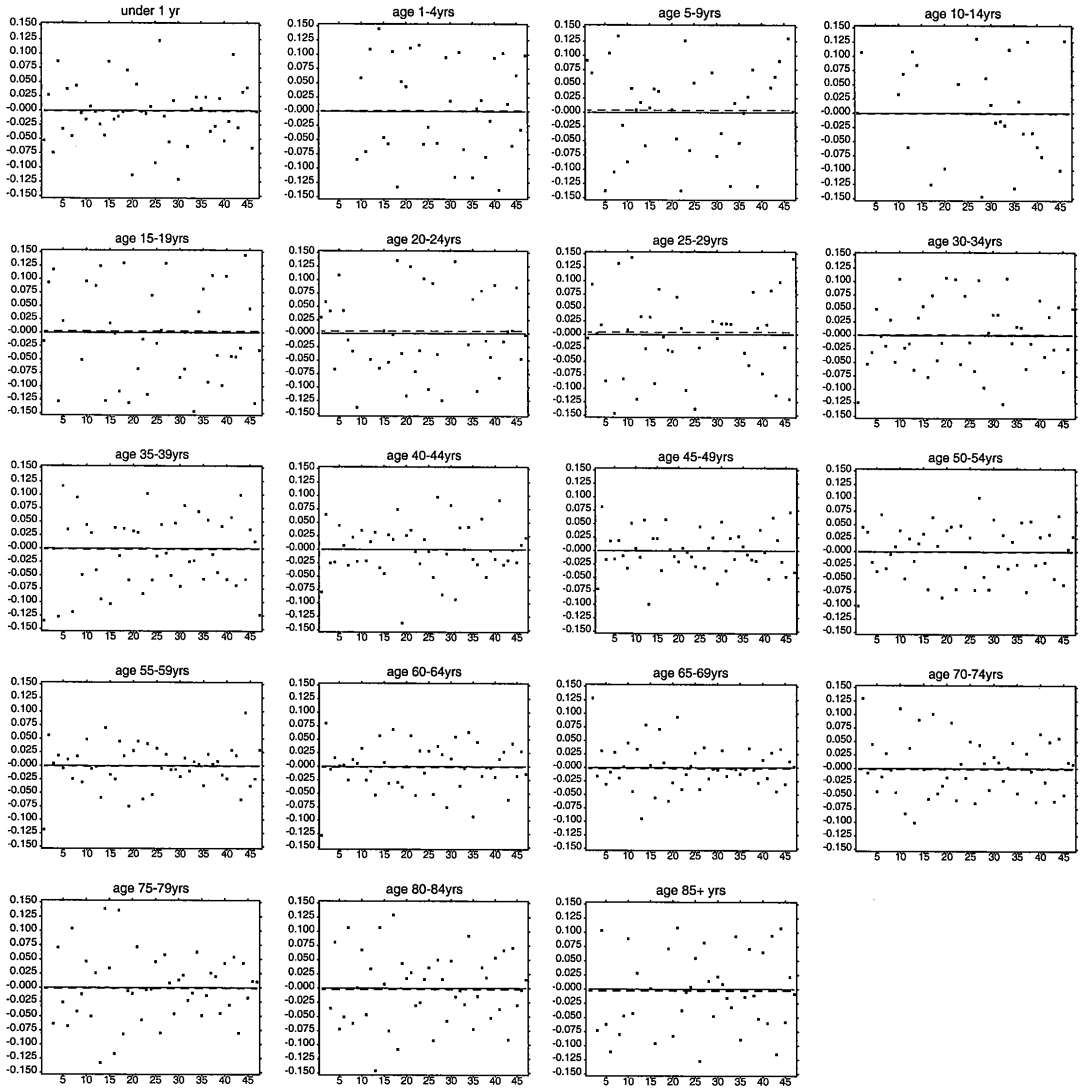
A.3 Male log crude mortality rate 2^{nd} differences (with mean), by year, for age



A.4 Female log crude mortality rates, by year, for age



A.5 Female log crude mortality rate 1st differences (with split means), by year, for age



A.6 Female log crude mortality rate 2^{nd} differences (with mean), by year, for age

Appendix B

(Some issues concerning prediction limits under the GLM approach)

Prediction limits for log(RF):

Prediction limits for

$\log RF_{GLM}(x, t_n + s) = \hat{\beta}_x s$, $s > 0$

are computed as

$\hat{\beta}_x s \pm z_{\alpha/2} se(\hat{\beta}_x)(s + t_n - t_0)$

where $\Phi(z_\alpha) = 1 - \alpha$ is the cumulative distribution function of N(0,1).

Approximate prediction limits for life expectancy at birth:

Consider the following age grouping

$[x, x + w_x)$; $x = x_0, x_1, ..., x_{k-1}$; $w_{x_i} = x_{i+1} - x_i > 0, i = 0,1,...,k-1, x_k = \omega$.

Focus on a specific time epoch t and write $m_{x,t} \equiv m_x$. Suppose x_{k-1} does not exceed 85 years, say, so that the approximation

$$q_x \cong \frac{w_x m_x}{(1 + f'_x w_x m_x)},\ x = x_0, x_1, ..., x_{k-2};\ q_{x_{k-1}} = 1 \qquad \text{(B.1)}$$

applies. Then the following steps are used to compute life expectancy, given $\{m_x\}$

$$l_{x+\omega_x} = l_x(1 - q_x),\ x = x_0, x_1, ..., x_{k-2} \qquad \text{(B.2)}$$

$$d_x = l_x - l_{x+\omega_x} = l_x \times q_x,\ x = x_0, x_1, ..., x_{k-2};\ d_{x_{k-1}} = l_{x_{k-1}} \qquad \text{(B.3)}$$

$$L_x = w_x(l_x - f'_x d_x),\ x = x_0, x_1, ..., x_{k-2};\ L_{x_{k-1}} = \frac{l_{x_{k-1}}}{m_{x_{k-1}}} \qquad \text{(B.4)}$$

with life expectancy at age x_i

$$e_{x_i} = \frac{T_{x_i}}{l_{x_i}}, \text{ where } T_{x_i} = \sum_{x=x_i}^{x_{k-1}} L_x \qquad \text{(B.5)}$$

and where the proportions f_x and $f'_x = 1 - f_x$ locate the data within each age cell $[x, x + w_x)$ (f_x being apportioned relative to x). Expression (B.2) implies

$$l_{x_i} = l_{x_0} \prod_{j=0}^{i-1} (1 - q_{x_j}),\ x = x_0, x_1, ..., x_{k-1},\ (\prod_{j=0}^{-1} = 0) \qquad \text{(B.6)}$$

and (B.3) and (B.4) imply

$$L_x = w_x(f_x l_x + f'_x l_{x+f_x}),\ x = x_0, x_1, ..., x_{k-2};\ L_{x_{k-1}} = \frac{l_{x_{k-1}}}{m_{x_{k-1}}}. \qquad \text{(B.7)}$$

Then from (B.5) and (B.7) we have that

$$T_{x_0} = \sum_{i=0}^{k-1} (w_{x_{i-1}} f'_{x_{i-1}} + w_{x_i} f_{x_i}) l_{x_i} \qquad \text{(B.8)}$$

subject to defining

$$w_{x_{-1}} f'_{x_{-1}} = 0,\ w_{x_{k-1}} f_{x_{k-1}} = \frac{1}{m_{x_{k-1}}}, \qquad \text{(B.9)}$$

so that (B.5), (B.8) and (B.6) imply that life expectancy at age x_0 is given as

$$e_{x_0} = \sum_{i=0}^{k-1} \{(w_{x_{i-1}} f'_{x_{i-1}} + w_{x_i} f_{x_i}) \prod_{j=0}^{i-1} (1 - q_{x_j})\} \tag{B.10}$$

where

$$1 - q_x = \frac{1 - w_x f_x m_x}{1 + w_x f'_x m_x} \tag{B.11}$$

from (B.1). Treating the forecast $\dot{e}_{x_0, t_n+s} \equiv e_{x_0}$ as a random variable, we seek to express the RHS of (B.10) as a linear function of the parameter estimates $\hat{\beta}_x$ by approximation. Introducing

$$\dot{m}_{x,t_n+s} = m_x = c_x \exp \beta_x s \cong c_x (1 + \beta_x s)$$

where $c_x = \hat{m}_{x,t_n}$, the most recently observed crude mortality rate at age x, we have that

$$1 - q_x \cong a'_x (1 - b_x \beta_x s) \tag{B.12}$$

on ignoring second and higher powers of β_x, where

$$a'_x = \frac{1 - w_x f_x c_x}{1 + w_x f'_x c_x}; \mathrm{b}_x = \frac{w_x c_x}{(1 - w_x f_x c_x)(1 + w_x f'_x c_x)}.$$

Finally, defining

$$x_0 = 0, a_i = \prod_{j=0}^{i} a'_{x_j}, a_{-1} = 1, b_i = \mathrm{b}_{x_i}$$

and re-instating the full notation, (B.10), together with (B.12) and (B.9), imply

$$\dot{e}_{0,t_n+s} \cong A + \sum_{i=0}^{k-1} B_i \hat{\beta}_{x_i} s$$

on ignoring second and higher powers of $\hat{\beta}_{x_i}$, where

$$A = \sum_{i=0}^{k-2} w_{x_i} (f_{x_i} a_{i-1} + f'_{x_i} a_i) + c_{x_{k-1}}^{-1} a_{k-2}$$

$$B_i = -b_i \{\sum_{j=i}^{k-3} a_j (w_{x_j} f'_{x_j} + w_{x_{j+1}} f_{x_{j+1}}) + a_{k-2} (w_{x_{k-2}} f'_{x_{k-2}} + c_{x_{k-1}}^{-1})\}, i = 0,1,...,k-3$$

$$B_{k-2} = -b_{k-2} a_{k-2} (w_{x_{k-2}} f'_{x_{k-2}} + c_{x_{k-1}}^{-1})$$

$$B_{k-1} = -a_{k-2} c_{x_{k-1}}^{-1}.$$

Thus, these results lead to

$$E(\dot{e}_{0,t_n+s}) \cong A + \{\sum_{i=0}^{k-1} B_i E(\hat{\beta}_{x_i})\} s$$

and

$$Var(\dot{e}_{0,t_n+s}) \cong \{\sum_{i=0}^{k-1} B_i^2 Var(\hat{\beta}_{x_i}) + 2 \sum_{all\ i,j;\ i>j} B_i B_j Cov(\hat{\beta}_{x_i} \hat{\beta}_{x_j})\} s^2.$$

The case $\{x_i\} = \{0,1,5,10,...,85\}, \omega = 120$, $\{w_{x_i}\} = \{1,4,5,5,...,5,35\}$, $i = 0,1,...,18$ and $k = 19$, in which we set $f_{x_0} = 0.15$ for males, $f_{x_0} = 0.16$ for females (Registrar General (1987): Appendix B), $f_{x_i} = 0.5 (i = 1,2,...,k-2)$, is of immediate interest.

List of Figures

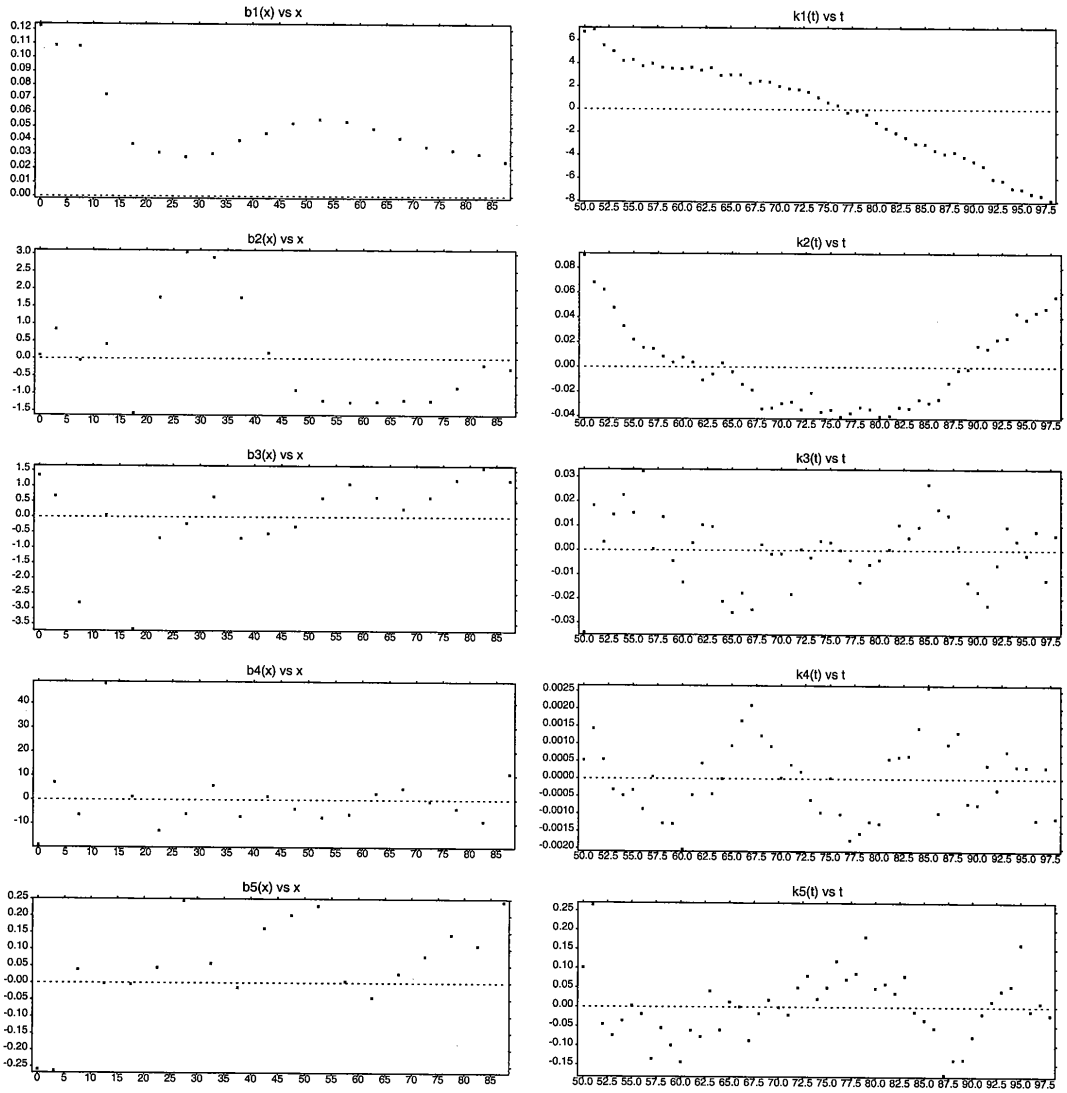


Fig 4.1a 1950-98 male experience, LC modelling, 1st five singular vectors

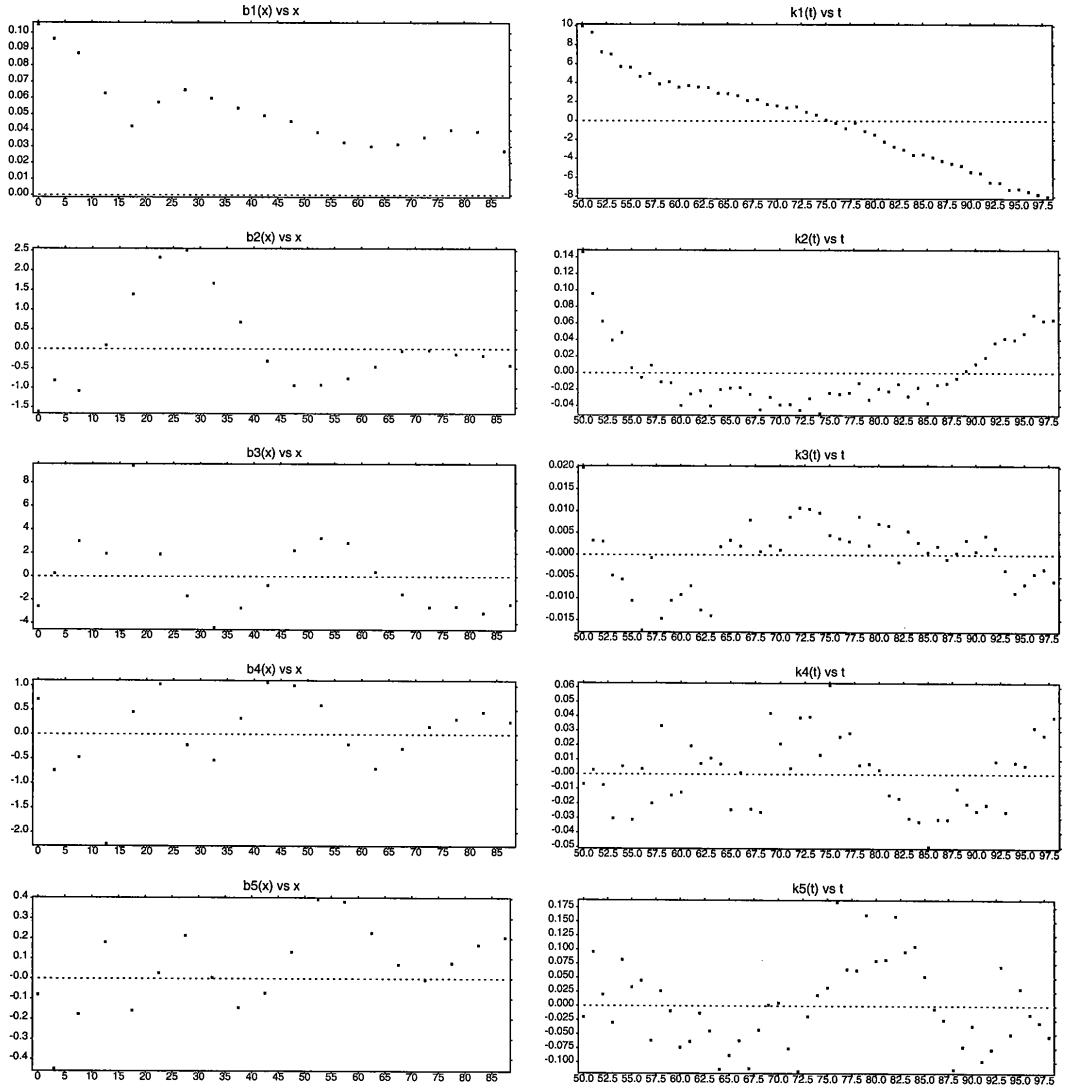


Fig 4.1b 1950-98 female experience, LC modelling, 1st five singular vectors

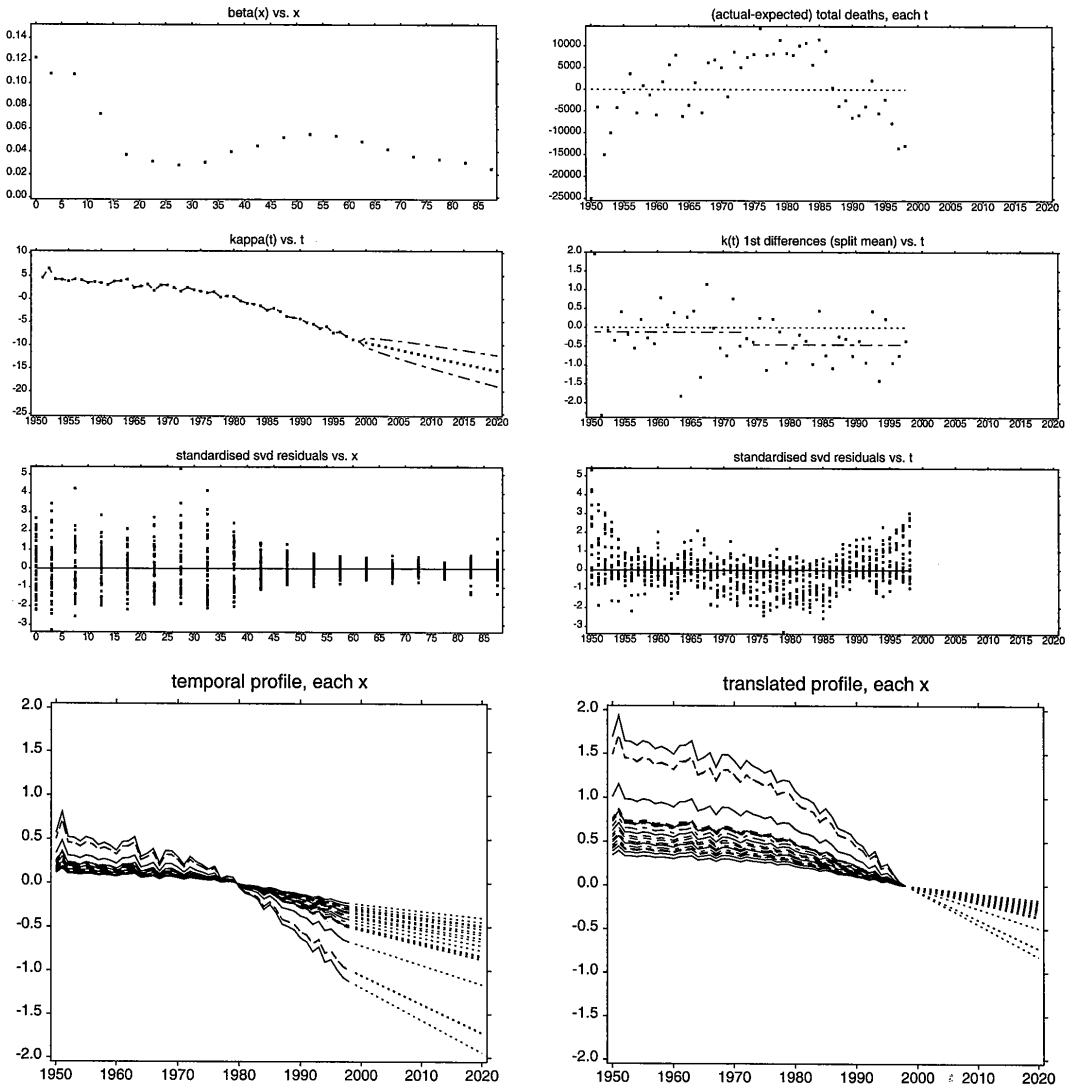


Fig 4.2a 1950-98 male experience, LC full base modelling

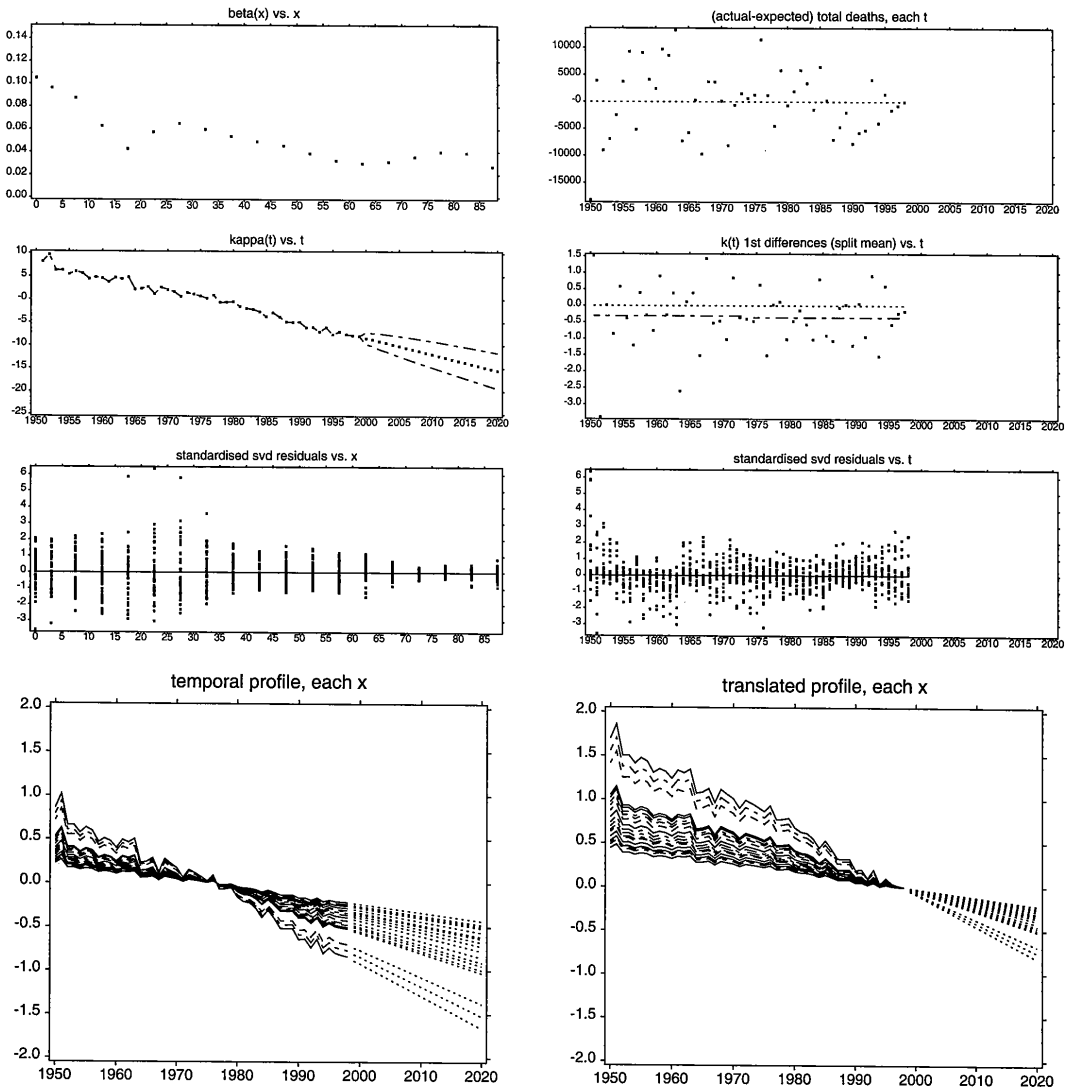


Fig 4.2b 1950-98 female experience, LC full base modelling

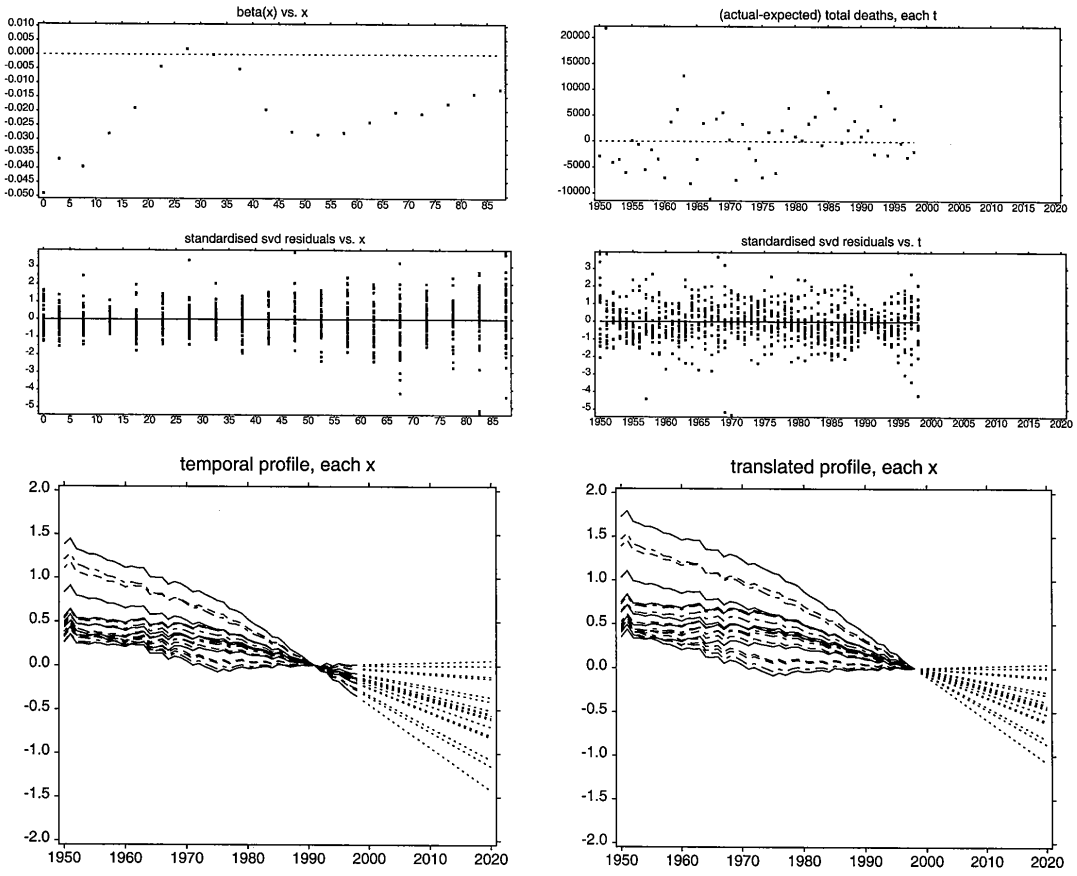


Fig 4.3a 1950-98 male experience, GLM 90-92 base modelling

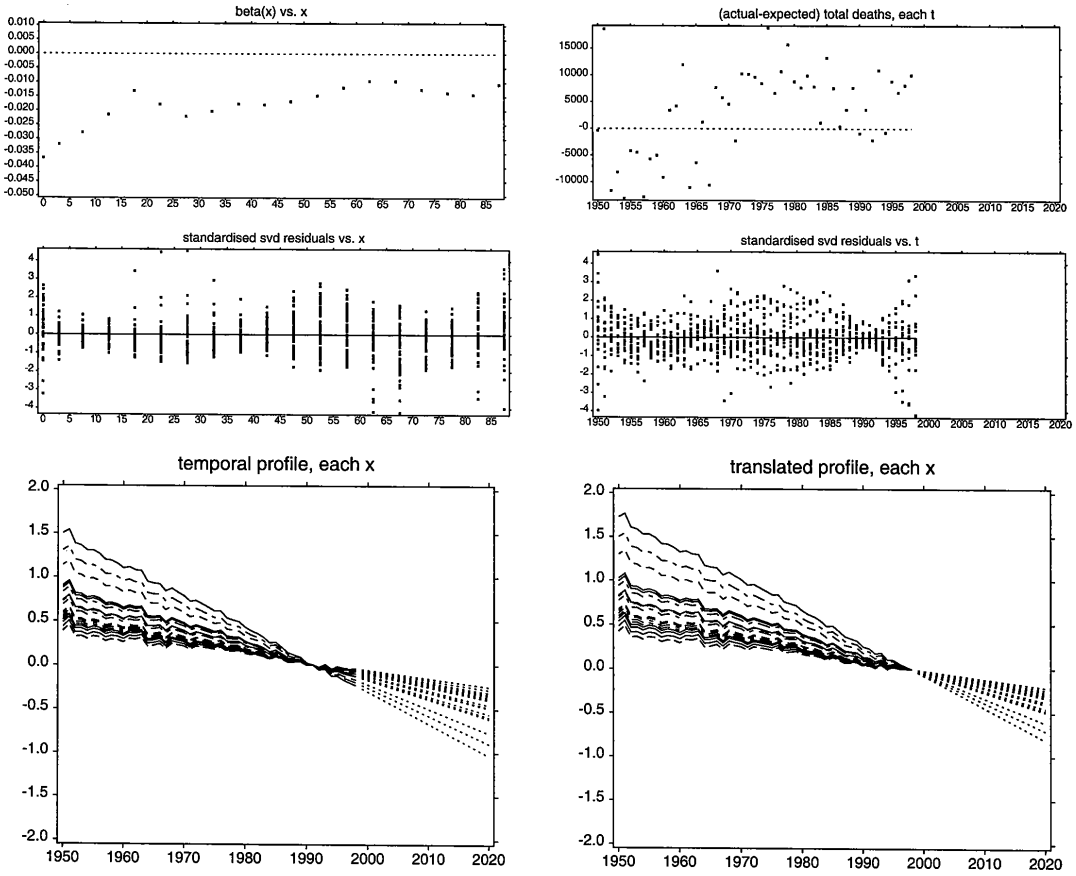


Fig 4.3b 1950-98 female experience, GLM 90-92 base modelling

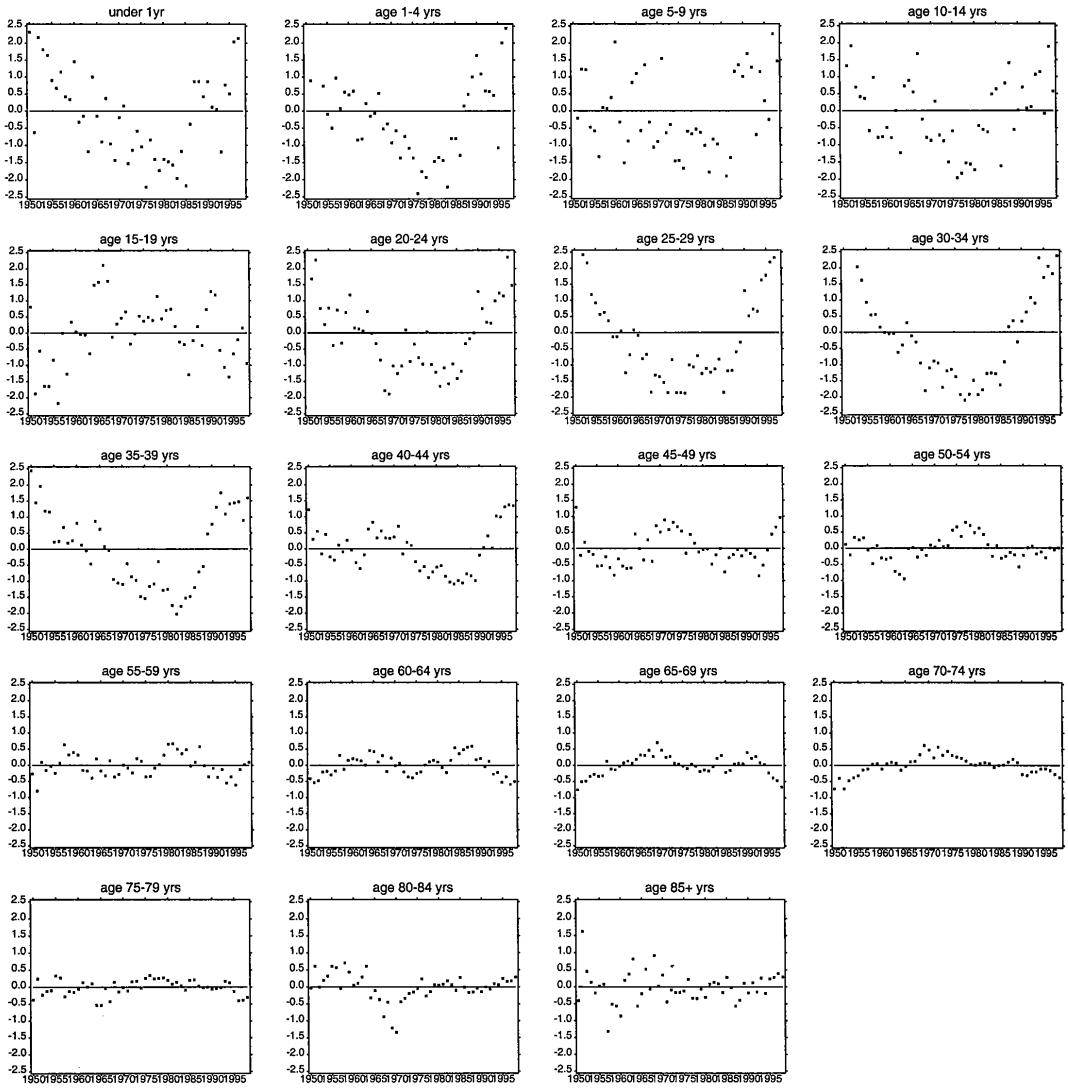


Fig 4.4a 1950-98 male experience, LC full base modelling, svd residuals

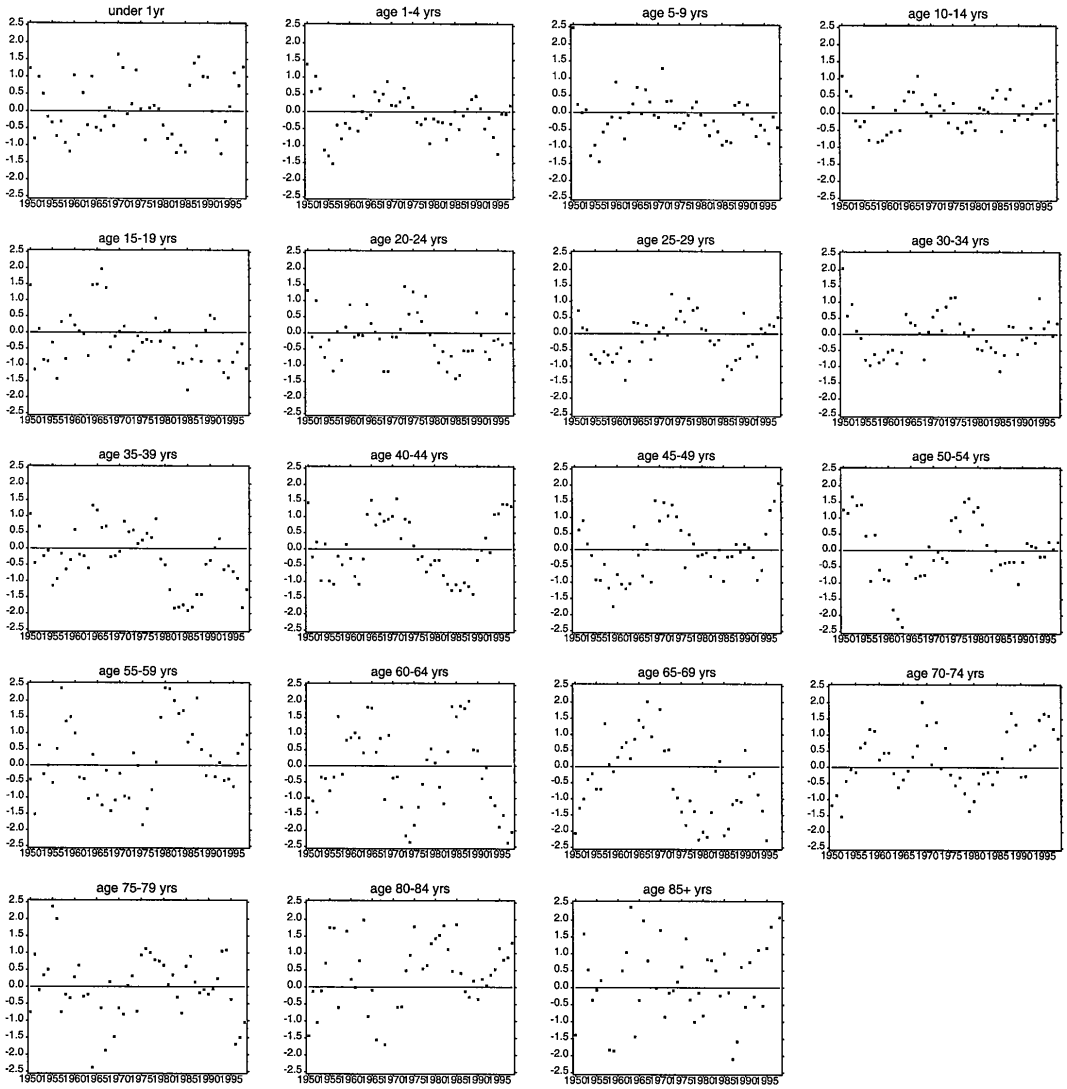


Fig 4.4b 1950-98 male experience, GLM 90-92 base modelling, svd residuals

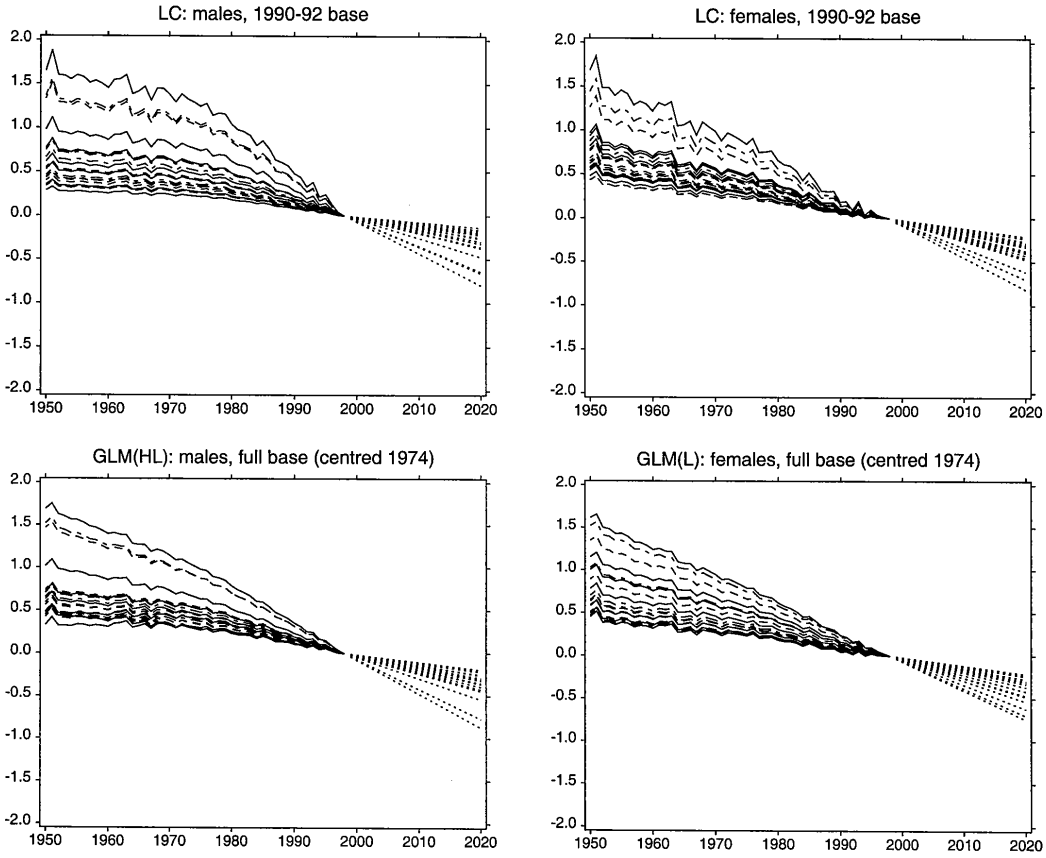


Fig 4.5 1950-98 experiences: LC 90-92 base and GLM full base modelling

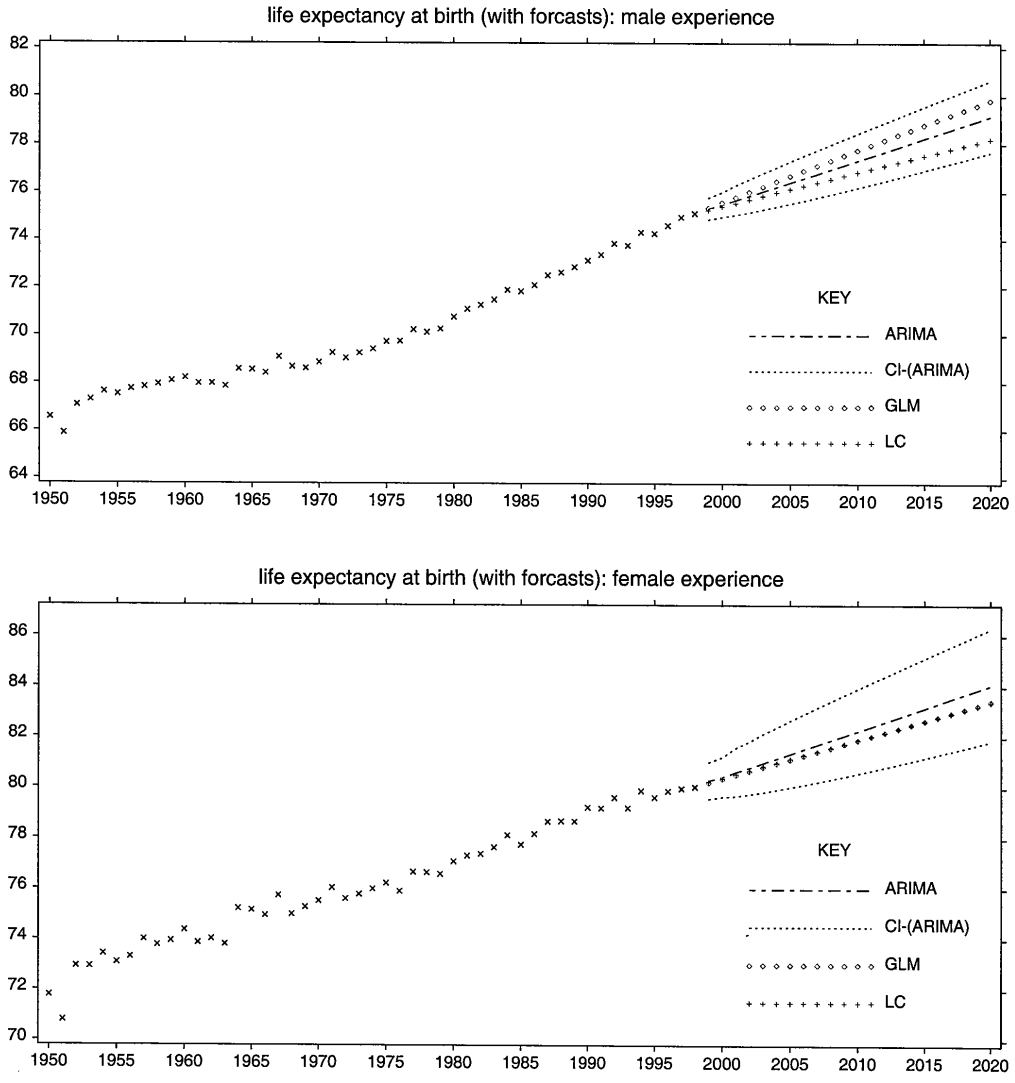


Fig 5.1 1950-98 male and female life expectancies with forecasts

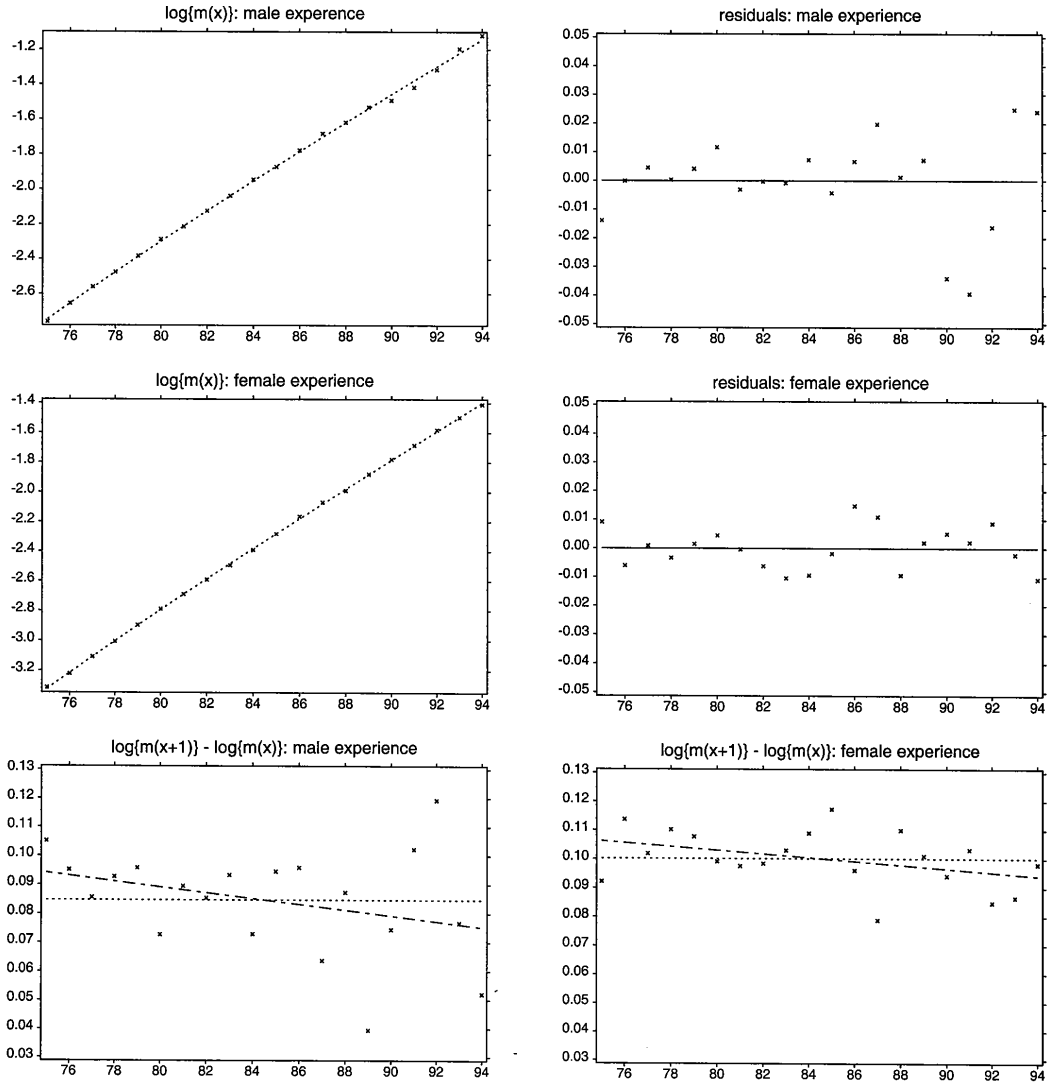


Fig 6.1 1990-92 experience by gender: analysis - log crude mortality rates by age

CITY City University London

DEPARTMENT OF ACTUARIAL SCIENCE AND STATISTICS

Actuarial Research Papers since 2000

---

122. Booth P.M. and Cooper D.R. The Tax Treatment of Pensions. April 2000. 36 pages. ISBN 1 901615 42 1

123. Walsh D.E.P. and Rickayzen B.D. A Model for Projecting the number of People who will require Long-Term Care in the Future. Part I: Data Considerations. July 2000. 37 pages. ISBN 1 901615 43 X

124. Rickayzen B.D. and Walsh D.E.P. A Model for Projecting the number of People who will require Long-Term Care in the Future. Part II: The Multiple State Model. July 2000. 27 pages. ISBN 1 901615 44 8

125. Walsh D.E.P. and Rickayzen B.D. A Model for Projecting the number of People who will require Long-Term Care in the Future. Part III: The Projected Numbers and The Funnel of Doubt. July 2000. 61 pages. ISBN 1 901615 45 6

126. Cooper D.R. Security for the Members of Defined Benefit Pension Schemes. July 2000. 23 pages. ISBN 1 901615 45 4

127. Renshaw A.E. and Haberman S. Modelling for mortality reduction factors. July 2000. 32 pages. ISBN 1 901615 47 2

128. Ballotta L. and Kyprianou A.E A note on the -quantile option. September 2000. ISBN 1 901615 49 9

129. Spreeuw J. Convex order and multistate life insurance contracts. December 2000. ISBN 1 901615 50 2

130. Spreeuw J. The Probationary Period as a Screening Device. December 2000. ISBN 1 901615 51 0

131. Owadally M.I. and Haberman S. Asset Valuation and the Dynamics of Pension Funding with Random Investment Returns. December 2000. ISBN 1 901615 52 9

132. Owadally M.I. and Haberman S. Asset Valuation and Amortization of Asset Gains and Losses in Defined Benefit Pension Plans. December 2000. ISBN 1 901615 53 7

133. Owadally M.I. and Haberman S. Efficient Amortization of Actuarial Gains/Losses and Optimal Funding in Pension Plans. December 2000. ISBN 1 901615 54 5

134. Ballotta L. -quantile Option in a Jump-Diffusion Economy. December 2000. ISBN 1 901615 55 3

135. Renshaw A. E. and Haberman S. On the Forecasting of Mortality Reduction Factors. February 2001. ISBN 1 901615 56 1

136. Haberman S., Butt Z. & Rickayzen B. D. Multiple State Models, Simulation and Insurer Insolvency. February 2001. 27 pages. ISBN 1 901615 57 X

137. Khorasanee M.Z. A Cash-Flow Approach to Pension Funding. September 2001. 34 pages. ISBN 1 901615 58 8

138. England P.D. Addendum to "Analytic and Bootstrap Estimates of Prediction Errors in Claims Reserving". November 2001. 17 pages. ISBN 1 901615 59 6

139. Verrall R.J. A Bayesian Generalised Linear Model for the Bornhuetter-Ferguson Method of Claims Reserving. November 2001. 10 pages. ISBN 1 901615 62 6

140. Renshaw A.E. and Haberman. S. Lee-Carter Mortality Forecasting, a Parallel GLM Approach, England and Wales Mortality Projections. January 2002. 38 pages. ISBN 1 901615 63 4

Statistical Research Papers

1. Sebastiani P. Some Results on the Derivatives of Matrix Functions. December 1995. 17 Pages. ISBN 1 874 770 83 2

2. Dawid A.P. and Sebastiani P. Coherent Criteria for Optimal Experimental Design. March 1996. 35 Pages. ISBN 1 874 770 86 7

3. Sebastiani P. and Wynn H.P. Maximum Entropy Sampling and Optimal Bayesian Experimental Design. March 1996. 22 Pages. ISBN 1 874 770 87 5

4. Sebastiani P. and Settimi R. A Note on D-optimal Designs for a Logistic Regression Model. May 1996. 12 Pages. ISBN 1 874 770 92 1

5. Sebastiani P. and Settimi R. First-order Optimal Designs for Non Linear Models. August 1996. 28 Pages. ISBN 1 874 770 95 6

6. Newby M. A Business Process Approach to Maintenance: Measurement, Decision and Control. September 1996. 12 Pages. ISBN 1 874 770 96 4

7. Newby M. Moments and Generating Functions for the Absorption Distribution and its Negative Binomial Analogue. September 1996. 16 Pages. ISBN 1 874 770 97 2

8. Cowell R.G. Mixture Reduction via Predictive Scores. November 1996. 17 Pages. ISBN 1 874 770 98 0

9. Sebastiani P. and Ramoni M. Robust Parameter Learning in Bayesian Networks with Missing Data. March 1997. 9 Pages. ISBN 1 901615 00 6

10. Newby M.J. and Coolen F.P.A. Guidelines for Corrective Replacement Based on Low Stochastic Structure Assumptions. March 1997. 9 Pages. ISBN 1 901615 01 4.

11. Newby M.J. Approximations for the Absorption Distribution and its Negative Binomial Analogue. March 1997. 6 Pages. ISBN 1 901615 02 2

12. Ramoni M. and Sebastiani P. The Use of Exogenous Knowledge to Learn Bayesian Networks from Incomplete Databases. June 1997. 11 Pages. ISBN 1 901615 10 3

13. Ramoni M. and Sebastiani P. Learning Bayesian Networks from Incomplete Databases. June 1997. 14 Pages. ISBN 1 901615 11 1

14. Sebastiani P. and Wynn H.P. Risk Based Optimal Designs. June 1997. 10 Pages. ISBN 1 901615 13 8

15. Cowell R. Sampling without Replacement in Junction Trees. June 1997. 10 Pages. ISBN 1 901615 14 6

16. Dagg R.A. and Newby M.J. Optimal Overhaul Intervals with Imperfect Inspection and Repair. July 1997. 11 Pages. ISBN 1 901615 15 4

17. Sebastiani P. and Wynn H.P. Bayesian Experimental Design and Shannon Information. October 1997. 11 Pages. ISBN 1 901615 17 0

18. Wolstenholme L.C. A Characterisation of Phase Type Distributions. November 1997. 11 Pages. ISBN 1 901615 18 9

19. Wolstenholme L.C. A Comparison of Models for Probability of Detection (POD) Curves. December 1997. 23 Pages. ISBN 1 901615 21 9

20. Cowell R.G. Parameter Learning from Incomplete Data Using Maximum Entropy I: Principles. February 1999. 19 Pages. ISBN 1 901615 37 5

21. Cowell R.G. Parameter Learning from Incomplete Data Using Maximum Entropy II: Application to Bayesian Networks. November 1999. 12 Pages ISBN 1 901615 40 5

22. Cowell R.G. FINEX : Forensic Identification by Network Expert Systems. March 2001. 10 pages. ISBN 1 901615 60X

23. Cowell R.G. When Learning Bayesian Networks from Data, using Conditional Independence Tests is Equivalant to a Scoring Metric. March 2001. 11 pages. ISBN 1 901615 61 8

Department of Actuarial Science and Statistics

Actuarial Research Club

The support of the corporate members

CGNU Assurance
Computer Sciences Corporation
Government Actuary's Department
HCM Consultants (UK) Ltd
KPMG
PricewaterhouseCoopers
Swiss Reinsurance
Watson Wyatt Partners

is gratefully acknowledged.

ISBN 1 901 615 63 4